

=> d que

L4 72 SEA FILE=HCAPLUS ABB=ON PLU=ON ("BARRY C"/AU OR "BARRY C C"/AU OR "BARRY C D"/AU OR "BARRY C DAVID"/AU OR "BARRY C E"/AU OR "BARRY C E III"/AU OR "BARRY C J"/AU OR "BARRY C L"/AU OR "BARRY C M"/AU OR "BARRY C M F"/AU OR "BARRY C MARVIN"/AU OR "BARRY C N"/AU OR "BARRY C P"/AU OR "BARRY C S"/AU OR "BARRY CAROLINE"/AU OR "BARRY CAROLYN B"/AU)

L5 126 SEA FILE=HCAPLUS ABB=ON PLU=ON ("BOUGUELERET L"/AU OR "BOUGUELERET LYDIE"/AU)

L6 130 SEA FILE=HCAPLUS ABB=ON PLU=ON ("CHUMAKOV I"/AU OR "CHUMAKOV I F"/AU OR "CHUMAKOV I M"/AU OR "CHUMAKOV I R"/AU OR "CHUMAKOV I S"/AU OR "CHUMAKOV I V"/AU OR "CHUMAKOV ILYA"/AU OR "CHUMAKOV ILYA M"/AU)

L7 12 SEA FILE=HCAPLUS ABB=ON PLU=ON (L*** OR L*** OR L***)

L8 27 SEA FILE=HCAPLUS ABB=ON PLU=ON (L4 AND (L5 OR L6 OR L7)) OR (L5 AND (L6 OR L7)) OR (L6 AND L7)

L9 302 SEA FILE=HCAPLUS ABB=ON PLU=ON (L4 OR L5 OR L6 OR L7)

L10 1 SEA FILE=HCAPLUS ABB=ON PLU=ON L9 AND BAP?

L11 23 SEA FILE=HCAPLUS ABB=ON PLU=ON L9 AND PROSTAT?

L12 37 SEA FILE=HCAPLUS ABB=ON PLU=ON L8 OR (L10 OR L11)

=> d ibib abs 1-37

L12 ANSWER 1 OF 37 HCAPLUS COPYRIGHT 2006 ACS on STN

ACCESSION NUMBER: 2006:367166 HCAPLUS
 DOCUMENT NUMBER: 144:405904
 TITLE: Polynucleotides encoding human TBC-1 protein,
 polymorphic markers therein, and uses thereof
 INVENTOR(S): Blumenfeld, Marta; Bougueleret, Lydie;
 Chumakov, Ilya
 PATENT ASSIGNEE(S): Genset S.A., Fr.
 SOURCE: U.S. Pat. Appl. Publ., 114 pp.
 CODEN: USXXCO
 DOCUMENT TYPE: Patent
 LANGUAGE: English
 FAMILY ACC. NUM. COUNT: 1
 PATENT INFORMATION:

PATENT NO.	KIND	DATE	APPLICATION NO.	DATE
US 2006084073	A1	20060420	US 2004-970750	20041020
PRIORITY APPLN. INFO.:			US 2004-970750	20041020

AB The invention concerns genomic and cDNA sequences of the human TBC-1 gene. The invention also concerns polypeptides encoded by the TBC-1 gene. The invention also deals with antibodies directed specifically against such polypeptides that are useful as diagnostic reagents. The invention further encompasses biallelic markers of the TBC-1 gene useful in genetic anal. The TBC-1 gene is located in a candidate region for prostate cancer on chromosome 4. Expression of TBC-1 gene leads to production of at least two mRNA mols. which have different first exons, exon 1 and 1bis, due to alternative splicing.

L12 ANSWER 2 OF 37 HCAPLUS COPYRIGHT 2006 ACS on STN

ACCESSION NUMBER: 2005:182920 HCAPLUS
 DOCUMENT NUMBER: 142:258503
 TITLE: Secreted polypeptide species in human plasma,
 detection assays for smaller proteins and tryptic
 peptides, and expression profiles useful for disease

OM protein - protein search, using sw model	GenCore version 5.1.9			
Run on:	June 29, 2006, 00:41:09 : search time 110.097 Seconds (without alignments)			
Scoring table:	BLOSUM62			
Searched:	Gapext 10.0 , Gapext 0.5			
Total number of hits satisfying chosen parameters:	2599679			
Minimum DB seq length:	0			
Maximum DB seq length:	200000000			
Post-processing:	Maximum Match 0% Maximum Match 100% Listing first 45 summaries			
Database :	A_GeneSeq_8;*			
	1: geneseqp1980s;*			
	2: geneseqp1990s;*			
	3: geneseqp2000s;*			
	4: geneseqp2001s;*			
	5: geneseqp2002s;*			
	6: geneseqp2003as;*			
	7: geneseqp2003bs;*			
	8: geneseqp2004s;*			
	9: geneseqp2005s;*			
	10: geneseqp2006s;*			
Pred.	No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.			
	SUMMARIES			
Result No.	Score	Query Match Length	DB ID	Description
1	8212	100.0	2144	AAB85029
2	8212	100.0	2144	ADQ88800
3	8204	99.9	2144	AEE02559
4	8195	99.8	2144	ADS4902
5	7635	93.0	2044	ADS34907
6	7627.5	92.9	1569	ABG1232
7	7626	88.5	2036	ADS34905
8	7265	88.5	2036	ADS34903
9	7266	88.5	2036	ADS34904
10	3694	45.0	734	ADQ6615
11	3694	45.0	734	AEC87545
12	3546.5	43.2	1229	ADQ6400
13	3102	37.8	624	AEE02884
14	2824.5	34.4	1149	ADE08012
15	2824.5	34.4	1149	ADU40382
16	1918	23.4	897	ADQ96220
17	1916	23.3	897	ADQ96152
18	1914	23.3	897	ADQ96218
19	1603	23.5	325	ADE09054
20	1603	19.5	325	ADU40573
21	1396.5	17.0	2096	ABE65342
22	1396.5	17.0	2096	ADQ9606
23	705.5	8.6	1798	ABJ26330

Post-processing: Maximum Match 0%
Maximum Match 100%
Listing first 45 summaries

RESULT 1

ID AAB85029 standard; protein: 2144 AA.
XX
AC AAB85029;
XX
DT 06-AUG-2001 (first entry)
XX
DE Protein encoded by BAP28 cDNA consisting of exons 1 to 45.
XX
KW BAP28; prostate; tumour; cancer; diagnostic; genetic analysis.
XX
OS Homo sapiens.
XX
FH Key location/Qualifiers
FT Misc-difference 1694 /label= Ser or Asn
FT Misc-difference 1854 /label= Ala or Val
FT Misc-difference 1967
FT Misc-difference /label= Asp or Asn
FT Misc-difference 2017 /label= Gly or Glu
PT XX
PT WO201006669-A2.
PT XX
PT PD 04-JAN-2001.
PT XX
PT PF 23-JUN-2000; 2000WO-IB001183.
PT XX
PT PR 25-JUN-1999; 90US-014132P.
PT PR 18-JAN-2000; 2000US-0176880P.
PT XX
PT PA (GST) SENSIT
PT XX
PT PI Barry C Bouquerel et L, Chumakov I, Cohen-Akenine A;
PT XX
PT DR WPI; 2001-367032/38.
PT DR N/P-SDB; AAF83909, AAF83910.
PT XX
PT New BAP28 polynucleotides and polypeptides overexpressed in prostate cancer cells for diagnosing prostate tumors, e.g. by hybridization or polymerase chain reaction assays.
PT XX
PT PS Claim 14; Page 297-304; 349pp; English.

Abj25730 Aspergill
Abj73779 Candida a
Abp11280 Human ova
Abp1280 Human
Abp2974 Proteins
Adl62658 Disease t
Aam17368 Peptide #
Abp36384 Peptide #
Aam29882 Peptide #
Abp31185 Peptide #
Abp21738 Protein #
Aam69542 Human bon
Aam57146 Human bra
Abp51215 Human liv
Aam5057 Peptide #
Abp31917 Human pep
Adk15818 Human ABC
Adk15814 Human ABC
Adl15793 Human ABC
Adp30671 Arabidops
Adp30670 Arabidops
Aag30663 Arabidops
Aag30669 Chicken A
Adf45530 Chicken A

The invention is directed to Bap22 polypeptides, Bap28 poly nucleotide sequences and regulatory region located at the 3' and 5' ends of the Bap28 coding region. The Bap28 polypeptides can be expressed by standard recombinant methodology. Bap28 polynucleotides and polypeptides have been found to be over expressed in prostate tumour cells, therefore levels of Bap28 expression and/or activity may be assayed (e.g. by polymerase chain reaction (PCR)) to diagnose patient suffering from or susceptible to prostate cancer. Antibodies specific for the Bap28 polypeptides are useful as diagnostic reagents. Biallelic markers for the Bap28 gene are useful in genetic analysis. The present sequence represents a protein encoded by a first cDNA sequence of the Bap28 gene consisting of the exons 1 to 45

QY	781	VAVEDSVFVFLFSLKKFTVALKAKPSKFPGKDIDWNPEOLKDSDRDYLHLHIGLFEMMLNCA	BAG
841	DAVEFRVLMKLFIKVHLEDFQFLPKFCSPVLTGSSLNPNSVKTUQTAIYGCAM	900	
841	DAVEFRVLMKLFIKVHLEDFQFLPKFCSPVLTGSSLNPNSVKTUQTAIYGCAM	900	
Db	901	LSSOKTQCKHOLASISPVTSLSNLGSNSPVKEVRRAOCOLAGIASPAPFILLDHLI	961
Db	901	LSOKTQCKHOLASISPVTSLSNLGSNSPVKEVRRAOCOLAGIASPAPFILLDHLI	961
QY	961	SKAEBITSDRAYVIDLATLFEELQREKELKSHKHOSETTKNLJSCVSCPSYAKDMK	961
Db	961	SKAEBITSDRAYVIDLATLFEELQREKELKSHKHOSETTKNLJSCVSCPSYAKDMK	961
QY	961	VLGQVNGEMVLQLSPLPMAPOLLEKIQKEPTAVLKDAMYLHTLGKYNEFVSULNEPK	1071
Db	961	VLGQVNGEMVLQLSPLPMAPOLLEKIQKEPTAVLKDAMYLHTLGKYNEFVSULNEPK	1071
QY	1021	VLGQVNGEMVLQLSPLPMAPOLLEKIQKEPTAVLKDAMYLHTLGKYNEFVSULNEPK	1071
Db	1021	VLGQVNGEMVLQLSPLPMAPOLLEKIQKEPTAVLKDAMYLHTLGKYNEFVSULNEPK	1071
QY	1081	SLDFRKAVHTTKELYAGMPTIQTALEKITKPRPAATSDEKVQQLRMLFDLWNCN	1141
Db	1081	SLDFRKAVHTTKELYAGMPTIQTALEKITKPRPAATSDEKVQQLRMLFDLWNCN	1141
QY	1141	SHCAOTVSSPFKGISVNAQAVRIELEPPDKAKPGLTVOQRQGMQOKQSODLESQEVG	1201
Db	1141	SHCAOTVSSPFKGISVNAQAVRIELEPPDKAKPGLTVOQRQGMQOKQSODLESQEVG	1201
QY	1201	GSYMDRVTLLLELOKKKLRSPOLIVPTFLNLSRCBPLPQFOGNNMYTKQILSCL	1201
Db	1201	GSYMDRVTLLLELOKKKLRSPOLIVPTFLNLSRCBPLPQFOGNNMYTKQILSCL	1201
QY	1261	NICQKISPOGGKIPKDLBREKENVELTYOCIRSEMPOPTHHALLLGTVAGIFPDKV	1321
Db	1261	NICQKISPOGGKIPKDLBREKENVELTYOCIRSEMPOPTHHALLLGTVAGIFPDKV	1321
QY	1321	HNMISFTENGANVRLDDTYSFQVINKTUMVTPALIOSDGSIESVERNEELVWKII	1321
Db	1321	HNMISFTENGANVRLDDTYSFQVINKTUMVTPALIOSDGSIESVERNEELVWKII	1321
QY	1381	SVFVDALPHVPEHRLPILVQLVPTGAEKFLWILLLFEQYVTTVLAAYGEKDAIL	1441
Db	1381	SVFVDALPHVPEHRLPILVQLVPTGAEKFLWILLLFEQYVTTVLAAYGEKDAIL	1441
QY	1441	BADTEFWFSVCCEFSVHQHQIQLSMLIQLYKLUPEKEETIPKAVSFNKSESOEMLQVF	1501
Db	1441	BADTEFWFSVCCEFSVHQHQIQLSMLIQLYKLUPEKEETIPKAVSFNKSESOEMLQVF	1501
QY	1501	NVETMSKQHRHKFTLSVNSNQLSNSNNLKKUVESGGPBILKGLERLETFVGYISA	1561
Db	1501	NVETMSKQHRHKFTLSVNSNQLSNSNNLKKUVESGGPBILKGLERLETFVGYISA	1561
QY	1561	VAQSMERNADKLTVKFWRALSKAYDLDKNUAIPLTETIPVTRGGLVGNPLPSYRKAL	1561
Db	1561	VAQSMERNADKLTVKFWRALSKAYDLDKNUAIPLTETIPVTRGGLVGNPLPSYRKAL	1561
QY	1621	DLNKNIQO 1629	1621
Db	1621	DLNKNIQO 1629	1621
RESULT 2			
ADQ89800		ADQ89800 standard; protein; 2144 AA.	
ID			
XX			
AC			
XX			
ADQ89800;			
DT			
21-OCT-2004 (first entry)			
XX			
DE			
Antagonist of cell cycle progression polypeptide #115.			
XX			
KW			
Cytostatic; cancer; cell division cycle; mitosis; meiosis; cell cycle progression.			
XX			
OS			
Homo sapiens.			

GenCore version 5.1.9
 Copyright (c) 1993 - 2006 Biocceleration Ltd.

OM protein - protein search, using sw model

Run on: June 29, 2006, 00:42:00 ; Search time 39.2894 Seconds

(without alignments) 3989.295 Million cell updates/sec

Title: US-09-603-665-5_COPY_1_1629
 Perfect score: 8212
 Sequence: 1 MTSLAQQLQRLLALPQSDASL.....NPLPSVRRKALDLNNKLQQ 1629
 Scoring table: BIOSUM62
 Gapop 10.0 , Gapext 0.5

Searched: 283416 seqs, 96216763 residues

Total number of hits satisfying chosen parameters: 283416

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
 Maximum Match 100%
 Listing first 45 summaries

Database : PIR 80:*

1: pir1:*

2: pir2:*

3: pir3:*

4: pir4:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No. Score Query Length DB ID

Description

Result No.	Score	Query	Length	DB ID	Description
1	673.5	8.2	1649	2 T39938	hypothetical protein SPBC23E6_04C - fission yeast (Schizosaccharomyces pombe)
2	654	8.0	1769	2 S53378	C;Species: Schizosaccharomyces pombe
3	536	6.5	1650	2 T27864	C;Date: 01-Dec-1999 #sequence_revision 03-Dec-1999 #text_change 09-Jul-2004
4	248	3.0	2570	2 T37919	C;Accession: T39938 R;Wood, V.; Ralandream, M.A.; Barrell, B.G.; Pohl, T. submitted to the EMBL Data Library, May 1998
5	241	2.9	2895	2 H83362	A;Reference number: Z21892
6	229	2.8	2059	2 T31933	A;Accession: T39938
7	228	2.8	3433	1 S28381	A;Status: Preliminary; translated from GB/EMBL/DDJB
8	228	2.8	4131	2 T21088	A;Molecule type: DNA
9	227.5	2.8	2954	2 T34156	A;Residues: 1-649 <HOO>
10	226	2.8	1447	2 R32909	A;Cross-references: UNIPROT:O50179; UNIPARC:UPI000013AF93; EMBL:AL023287; PIDN:CA18872.1
11	226	2.8	1830	2 E83909	A;Experimental source: strain 972b+; cosmid c23E6
12	224	2.7	2297	2 A32494	A;Genetics:
13	219.5	2.7	2819	2 A00551	A;Gene: SPBC23E6_04C
14	215	2.6	2663	1 S28261	A;Map position: 2
15	212.5	2.6	4385	2 T29042	Query Match 8.2%; Score 673.5; DB 2; Length 1649; Best Local Similarity 20.5%; Pred. No. 1.7e-25; Matches 348; Conservative 263; Mismatches 490; Indels 53; Gaps 61;
16	212	2.6	2108	2 S28473	Db
17	210.5	2.6	2269	2 T28677	Qy 2 TSLAQQLQRLLALPQSDASLRSDEV--ASLIFDRKEAATIDRDTAFAIGCTGLEELGI 58
18	209	2.5	2401	2 T28676	Db 3 SSIQOKLNII--QSN-NVLKINKIRRAPSILYDPKVAADMDELEIYTAVSGFHFELAVH 58
19	208	2.5	4717	2 T15181	Qy 59 DPSPEOFAPPLFSQAKTLERSVQTKAVNKOLDENLSFLIHLSPYFELKPAQCKLWLI 118
20	207.5	2.5	1655	2 S47446	Db 59 ERLLYFEKTILQEPOSYVQDYLTRINRNEKIDLCVOLILAPFTETKNAIKVLEWI 118
21	207.5	2.5	4859	2 S71713	Qy 119 HPHTHLYNQOSLIAACVLPYVHTRIVFVRLKINNSKHMFWLIPVKOSCVPLAKGL 178
22	207	2.5	3744	2 S28473	Db 119 RRSIHEVYVSDFBILSRLPFDHPFARILGSK-PKSRPFLLENIAKMP-VTISRADI 176
23	206	2.5	1919	2 T40032	Qy 179 ITCYKOLGFMDIFCISLVTKSYKVEAYPGSSAQLRVLTAVASTIVSALVAAEDVSND- 237
24	204	2.5	34734	2 S47334	Db 177 VHALSDRKEFFAMFAQFVNQTAESHNNP-----BLARFWAGTMVEVLVWHSSNEDP 229
25	202.5	2.5	3493	2 320 NYILDREFLRYSVAVSVYSSI-DPQIAGFMILSTIASLPLSPSIIPPVLVSAITDRLSP- 287	
26	202	2.5	3079	1 RGSY12	Qy 294 PSILKDGSLCLIVLU-----ORQKESLG----- 317
27	201.5	2.5	2672	2 A48126	Db 288 -DNKPALICVGHLQPCSSSEFDHOLEKRSFGASSLILBSQEBRHLDERFVSYWNL 346
28	2.4	1957	2 T38077	Qy 318 -----KCPFPHCNCVPDLITIHGISBTYDPSLPLRMLPHLVSIIHHVTCGETEGMD 372	
29	199.5	2 1964	2 A59282	Db 347 IKSRKQDKKRLISLD-TSISQIRVTHEQAKPLSVIP-----VNQDPKAL-- 392	
					Qy 373 QTVKRHLLEAILTKISLKNQNLPHLLASLIFEEVITYSSSOBENDSNKVSSLNQFPLPLIRL 432

ALIGNMENTS

RESULT 1

T39938
 hypothetical protein SPBC23E6_04C - fission yeast (Schizosaccharomyces pombe)
 C;Species: Schizosaccharomyces pombe
 C;Date: 01-Dec-1999 #sequence_revision 03-Dec-1999 #text_change 09-Jul-2004
 C;Accession: T39938
 R;Wood, V.; Ralandream, M.A.; Barrell, B.G.; Pohl, T.
 submitted to the EMBL Data Library, May 1998
 A;Reference number: Z21892
 A;Accession: T39938
 A;Status: Preliminary; translated from GB/EMBL/DDJB
 A;Molecule type: DNA
 A;Residues: 1-649 <HOO>
 A;Cross-references: UNIPROT:O50179; UNIPARC:UPI000013AF93; EMBL:AL023287; PIDN:CA18872.1
 A;Experimental source: strain 972b+; cosmid c23E6
 C;Genetics:
 A;Gene: SPBC23E6_04C
 A;Map position: 2

Query Match 8.2%; Score 673.5; DB 2; Length 1649;
 Best Local Similarity 20.5%; Pred. No. 1.7e-25;
 Matches 348; Conservative 263; Mismatches 490; Indels 53; Gaps 61;

Qy 2 TSLAQQLQRLLALPQSDASLRSDEV--ASLIFDRKEAATIDRDTAFAIGCTGLEELGI 58

Db 3 SSIQOKLNII--QSN-NVLKINKIRRAPSILYDPKVAADMDELEIYTAVSGFHFELAVH 58

Qy 59 DPSPEOFAPPLFSQAKTLERSVQTKAVNKOLDENLSFLIHLSPYFELKPAQCKLWLI 118

Db 59 ERLLYFEKTILQEPOSYVQDYLTRINRNEKIDLCVOLILAPFTETKNAIKVLEWI 118

Qy 119 HPHTHLYNQOSLIAACVLPYVHTRIVFVRLKINNSKHMFWLIPVKOSCVPLAKGL 178

Db 119 RRSIHEVYVSDFBILSRLPFDHPFARILGSK-PKSRPFLLENIAKMP-VTISRADI 176

Qy 179 ITCYKOLGFMDIFCISLVTKSYKVEAYPGSSAQLRVLTAVASTIVSALVAAEDVSND- 237

Db 177 VHALSDRKEFFAMFAQFVNQTAESHNNP-----BLARFWAGTMVEVLVWHSSNEDP 229

Qy 294 PSILKDGSLCLIVLU-----ORQKESLG----- 317

Db 288 -DNKPALICVGHLQPCSSSEFDHOLEKRSFGASSLILBSQEBRHLDERFVSYWNL 346

Qy 318 -----KCPFPHCNCVPDLITIHGISBTYDPSLPLRMLPHLVSIIHHVTCGETEGMD 372

Db 347 IKSRKQDKKRLISLD-TSISQIRVTHEQAKPLSVIP-----VNQDPKAL-- 392

Qy 373 QTVKRHLLEAILTKISLKNQNLPHLLASLIFEEVITYSSSOBENDSNKVSSLNQFPLPLIRL 432

Db 393 QSYRFLDSLVIQPERKEKGKDNLNITL-----ODKKKSSTFSKKDREVL---- 436
 Qy 433 ESKYPRTLDVVLEEHLEKETADLKQELPHOFVSUSTSGKYQFLADSITSLMSLNPLA 492.
 Db 437 -----LKKSETDQSFTSBOCLAYADSA----ADLSVVFISL---- 470
 Qy 493 PVRILLAMHHLKKIMKTSKEGVNDSSFIKEAVLARGDDNIDVLSAISAFEFKEFSEV 552
 Db 471 -----LSFG-DKPELFCA-----NGSERI 492
 Qy 553 TISNLINIFORAEISKNGEWYEVKIAADILIKEELISENDQNSNQVVCLPFFV--IN 610
 Db 493 IILSIEL-----RTKIEENKDWDYQII--LPVNLVLSQ 524

Qy 611 NDDTESEAEMKIAILYLSKGICSLHPLLRGWEALENVIKSTKKGKLIGVANGKMIELAD 670
 Db 525 SKDEEVRSRALNLLT-----FLELRNE 547
 Qy 671 NINLGDPSSMLKNEVEDLISVGEEBESPNLKQKTFPHVILSVLVCSSLIKETHPPFAIRVF 730
 Db 548 NL-----EFSIYGMDNDN----- 562
 Qy 731 SLLQQKIKKLESITAVIPSEHNMEDRGPFLVPELNAHYVEELNSTORVAVEDSFLV 790
 Db 563 ---KOLRWLSPET---KYYCSDLJLDRSSEIGL-----DGTYLF 596
 Qy 791 FSLKFKFYALKAPSKFKPGDIWNPQLKEDSRDYLHLLIGLFEMMLNGADAVHFRVLMK 850
 Db 597 SIVPERLUTERKERNASK-----EIAVT 619
 Db 851 LFKVHHLDFVQFLKFCSVLWTVGSSLNPLNCVKVTLQTOALYVFCAMLSQTKOCH 910
 Db 620 SFLSH-----AACSKLN----- 633
 Qy 911 QLASSISSPVVTSLLINLGSPVKEVRRAIOCLQMSVASYPFLVLLDHLISREITSDA 970
 Db 634 -----VTRVILLEILTRVHG-----KVBD---- 651
 Qy 971 AYVTDQDLATLFEELQREKKLKSQHQLSETLKNLUSCVCSPVAKOLMKVITQGVNGEMV 1030
 Db 652 ---AKQKILLPPLQESEFS-EKFKVSKRETEVAUNCFNHTS--FTSISLSESSNIV 704
 Qy 1031 LSQOLPMAEOLAKIQKOEPTAVLKDTEAMVHLTIGKYNFESVSLNEDPKSLDIFKRAH 1090
 Db 705 LSCAI-CRRIVE-IOS-----HL-----KDPORL-EFKAVI 733
 Qy 1091 TYKELYAGMPTQITAKERKTFEEAISDEKVOQKULMLEDLVNCKNSICAOVSS 1150
 Db 734 SODE---OPHYVVDVDSIKP-----DTV 755
 Qy 1151 FKGSISVNAEQRVTELEPPDKAKLGTVOOKRROKMQKKSQDLSVQEVGGSYWORYTL 1210
 Db 756 FK-----KLGSVVLEKPKPATRKRD-----HIFDSDYTRI 794
 Qy 1211 LEIJLQHKKKLRSPOQILVPLTNLCSRCLPLOEGNNYTKOLLSCLLN----CQK 1265
 Db 795 UEBETKNAASVPLKASF-LFEVILNSVI-ALKEDIVSSNLYQOLLGIVENIGASPIE 852
 Qy 1266 LSFDGGKIPKPDIDEEKENVELLWOCIRLSEMPOTHHALLIGTVAIGIPDKVLRNIMS 1325
 Db 853 LSP-----SIRDITLGCRSTINNPFQIONRNLALLVSLANAAPEAVLHGMP 899
 Qy 1326 ITPFMGAMVRLDDTYSQVINKTOMTIALQ--SPSGDSTEVSERNVEETWKITSVF 1383
 Db 900 ITPFMGSTVLSRDRDAAFSHVIOVKWVISALLRGKDFDSS-----LYSCF 947
 Qy 1384 VDALPHVPEHRLFLVOLVDTGAEKELWILLLFFQVYKTVLAAYGKAILEAD 1443
 Db 948 VNAFPFTPQHRRLRLYRVLQQTGSNRLSVLUU----QFAEMILLA----KSTNTVVAI 998
 Qy 1444 TEFWSVSCFVSQHQQLQISLMNTIQLYLLKPEK--BETIPKAWFSNKESOBEMLOQFN 1501

Db 999 HDPCLTIVQSFVADRGISINQCSRCFLKSLSJEQNSDSNQSKAVSLIKLDELP----MD 1053
 Qy 1502 VEPHTSKOLRIPKFLSVFMMQOLLSSNNF---LKKVYESCGPEILKGLEERULETVLGYI 1558
 Db 1054 VOLATLGLSLR-VKVLRL--ISLVSKAKNPAZDLAKIMENSIDSFVR-IQAGLFESKLLI 1109
 Qy 1559 SAVAQS---MERNADKLTVKWRALLSKAVDLDKVNALIPTETFIPVIRGLVGNPLPSV 1615
 Db 1110 TUHQSSNEME-----LGHVYVALRSVTHLNLFCITVLGLGLHDERRALL 1155
 Qy 1616 RKKALDLANKLQQ 1629
 Db 1156 REKALSTVQQRVQ 1169

RESULT 2

S53378 probable membrane protein YJL109c - Yeast (*Saccharomyces cerevisiae*)

N;Alternate names: hypothetical protein J0808

C;Species: *Saccharomyces cerevisiae*

O;Date: 05-May-1995 #Sequence_revision 01-Sep-1995 #text_change 09-Jul-2004

C;Accession: S53378;

R;Rasmussen, S.W.

A;Submitted to the Protein Data Library, February 1995

A;Description: A 37.5 kb region of yeast chromosome X includes the SME1, MEF2, GSH1 and C

C;Reference number: S53378;

R;Accession: S53378;

A;Molecule type: DNA

A;Cross-references: UNIPROT:P42945; UNIPARC:UPI0000052F40; EMBL:X85021; NID:9728698; PID:

A;Reference number: S56876

A;Accession: S56887

A;Molecule type: DNA

A;Cross-references: UNIPARC:UPI0000052F40; EMBL:Z49384; NID:91008292; PID:91008293; MIPS:

R;Rasmussen, S.W.

Yeast 11, 873-883, 1995

A;Title: A 37.5 kb region of yeast chromosome X includes the SME1, MEF2, GSH1 and C

A;Reference number: S57357;

R;Accession: S57357

A;Status: nucleic acid sequence not shown; translation not shown

A;Molecule type: DNA

A;Cross-references: 1-1769 <RAF>

A;Cross-references: UNIPARC:UPI0000052F40; EMBL:X85021; NID:9728698; PID:CAA59385; PII

A;Note: the nucleotide sequence was submitted to the EMBL Data Library, February 1995

C;Generics:

A;Cross-references: SGD:S0003645

A;Map position: 10L

C;Keywords: transmembrane protein

Query Match 8.0%; Score 654; DB 2; Length 1769;
 Best Local Similarity 20.8%; Pred. No. 1.7e-24; Gaps 65;
 Matches 366; Conservative 252; Mismatches 542; Indels 598; Gaps 65;
 Qy 1 1 MTSIADQOLRIPALPQDASL---ISRDEVAISLFLDPKEMATIDRTAFATIGCTGELL 56
 Db 1 MSSISDQOLQASNNATVALDKRKRQKLHSASLINSKTATQDQDFIFENASKALEELS 60
 Qy 57 GIDPSDEQFAPLFLQAKTERSVOTKVNKOLDENISFLHLSPPYFLKPAQCKLE 116
 Db 61 QTEPKFAIFSRSLFSESSISLDRNVQPKKEEJTKDQDNAYAFLASSKWWIAPTLATEW 120
 Qy 117 LIHRPHIILYQDLSIACVLPYHETRIFVRYIQLKINNSGRHWFLP-----VK 167
 Db 121 LYRFOHVKVTEMUMLISTUNNYQTFVKRILSIK-----LPPFLNCJSNFVR 169
 Qy 168 OSGVPVLAQKTVJTHCVDLGMDFICSLVTVSKVPAEYPOSSAQRLVLAFASTIVS 227
 Db 170 SEKPPITAL-TM-KLNDMDPLKLYSVDLQCGKATYHQ-----LIFTCCFIN- 219
 Qy 228 LVAABDVSDN1IAKLFPPYI---QKGKSSLPDYRATYMICQISVKTMENTVNSLA 283

Copyright (c) 1993 - 2006 Biocceleration Ltd.

GenCore version 5.1.9

Search time 166.656 Seconds
(without alignments)

941.668 Million cell updates/sec

OM protein - protein search, using sw model

Run on: June 29, 2006, 00:41:34 ; Search time 166.656 Seconds

40 600 7.3 120 2 Q6P664_HUMAN

Title: US-09-603-665-5_COPY_1_1629

Perfect score: 8212

Sequence: 1 MTSLAQDQLALPQSDASL.....NPPLPSVERRKALDILNNKLQQ 1629

Scoring table: BLOSUM62

Searched: Gapop 10.0 ; Gapext 0.5

Total number of hits satisfying chosen parameters: 2849598

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Maximum Match 0%

Listing First 45 summaries

Database : UniProt 7.2;*

1: uniprot_sprot:*

2: uniprot_trembl:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match length	DB ID	Description
1	8204	99.9	2144	1 HEATR_HUMAN
2	7755.5	94.4	2063	2 Q5T3Q7_HUMAN
3	4891	59.9	1180	2 Q3UNW7_MOUSE
4	4515	55.0	1090	2 Q3TNB8_MOUSE
5	4118	50.1	2159	2 Q7SY48_BRARE
6	4072	49.6	984	2 Q3TMK0_MOUSE
7	3589	43.7	2288	2 Q4T7Z5_TETNG
8	3099	37.7	733	2 Q8BLJ4_MOUSE
9	3095	37.7	733	2 Q3T9I0_MOUSE
10	2973	36.2	1106	2 Q6P197_HUMAN
11	2298	1278	2 Q7T152_BRARE	
12	2298	1336	2 Q7T152_BRARE	
13	2181.5	26.6	958	1 HEATR_MACFA
14	1916	23.3	897	2 Q8NTL7_HUMAN
15	1444	17.6	2104	2 Q7PWD5_ANOGA
16	1398.5	17.0	2096	2 Q6AWS2_DROME
17	1396.5	17.0	2096	1 HEATR_DROME
18	948.5	11.6	743	2 Q3V1X6_MOUSE
19	852	10.4	168	2 Q2K1F2_BOVIN
20	832	10.1	2251	2 Q4P937_USTMA
21	809.5	9.9	2	Q8T9E7_DROME
22	799.5	9.7	2237	2 Q54MLA_DICDI
23	786.5	9.6	1830	1 HEATR_ARATH
24	774	9.4	2021	2 Q5KBQ2_CRYNE
25	767	9.3	2021	2 Q5SNBT_CRYNE
26	763.5	9.3	1857	2 Q6BXQ5_DEBHA
27	763.5	9.2	1801	2 Q5B1X5_EMET
28	757.5	9.2	2122	2 Q5CAF8_ORYSA
29	715.5	8.7	1802	2 Q2ULC6_ASPAR
30	707.5	8.6	1770	2 Q6FT93_CANGA
31	707.5	8.6	2	Q4WL19_ASPPU

Scored: No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

RESULT¹
HEATR_HUMAN STANDARD; PRT, 2144 AA.
ID HEATR_HUMAN
AC Q9H83; Q5T3Q8; Q9WW23;
DT 01-JUN-2001 integrated into UniProtKB/Swiss-Prot.
DT 19-JUL-2005 sequence version 3.
DE HEAT ISOPET-containing Protein 1 (Protein BAP28).
Name=HEATR1; Synonyms=BAP28;
OS Homo sapiens (Human)
OC Eukaryote; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini; Hominidae;
OC Homo
RN NCBI_TaxID=9606;
RP GLY-2017.
RA Bougueleret L, Chumakov S, Barry C, Cohen-Akenine A.;
RT "A novel BAP28 gene and protein."
RL Patent number WO2005047665, 04 JAN 2001.
RN [2]
RP NUCLEOTIDE SEQUENCE AND VARIANTS SER-1694; ALA-1854; ASP-1967 AND
NUCLEOTIDE SEQUENCE [LARGE SCALE GENOMIC DNA].
RL Human chromosome 1, international sequencing consortium:
RN Submitted (MAY-2005) to the EMBL/GenBank/DBJ databases.
[3]
RP NUCLEOTIDE SEQUENCE [LARGE SCALE MRNA] OF 1777-2144.
Pubmed=14702039; DOI=10.1038/ng1285;
Ota T., Suzuki Y., Nishikawa T., Otsuki T., Sugiyama T., Irie R.,
Wakamatsu A., Hayashi K., Sato H., Nagai K., Kimura K., Makita H.,
Sekine M., Obayashi M., Nishi T., Shibahara T., Tanaka T., Ishii S.,
Yamamoto J., Saito K., Kawai Y., Isono Y., Nakamura Y., Nagahari K.,
Murakami K., Sudo H., Hosoi T., Koku T., Kodaira H., Kondo H., Sugawara M.,
Takahashi M., Kanda K., Yokoi T., Furuya T., Kikkawa E., Omura Y.,
Abe K., Kamihara K., Katsuta N., Sato K., Tanikawa M., Yanazaki M.,
Niromiya K., Ishibashi T., Yamashita H., Murakawa K., Fujimori K.,
Tani H., Kimuchi M., Watanabe M., Hirakawa S., Chiba Y., Ishida S.,
Ono Y., Takiguchi S., Watanabe M., Yosida M., Hotuda T., Kusano J.,
Kanehori K., Takahashi-Fujii A., Hara H., Tanase T.-O., Nomura Y.,
Toyagi S., Komai F., Hara T., Takeuchi K., Arita M., Imose N.,
Mutashiro K., Yuuki H., Oshima A., Sasaki N., Aotsuka S.,
Yoshikawa Y., Matsunawa H., Ichihara T., Shiohata N., Sano S.,
Moriya S., Momiyama H., Satoh N., Takami S., Terashima Y., Suzuki O.,
Nagawa S., Senoh A., Mizoguchi H., Goto K., Shimizu F., Wakabe H.,
Hishigaki H., Watanabe T., Sugiyama A., Takemoto M., Kawakami B.,
Yamazaki M., Watanabe K., Kumagai A., Ittakura S., Fukuzumi T.,
Fujimori Y., Komiyama M., Tashiro H., Tanigami A., Fujiwara T.,
Ono T., Yamada K., Fuji Y., Ozaki K., Hirao M., Ohmori Y.,
Kawabata A., Hikiji T., Kobatake H., Inagaki H., Ikeda Y., Okamoto S.,
Okitani R., Kawakami T., Neguchi S., Itoh T., Shigeta K., Semba T.,
Matsumura K., Nakajima Y., Mizuno T., Morinaga M., Sasaki M.,
Togashi T., Oyama M., Hata H., Watanabe M., Komatsu T.,
Mizushima-Sugano J., Satoh T., Shirai Y., Takahashi Y., Nakagawa K.,
Okumura K., Nagase T., Nomura N., Kikuchi H., Masuho Y., Yamashita R.,

07rz08 neurospora
06C157 kluveromyces
059246 candida alb
075418 asbya goss
06179 schizosacch
P42945 saccharomyces
06C517 yarrowia li
Q41696 gibberella
Q6P664 homo sapien
Q61v68 caenorhabdi
Q23495 caenorhabdi
Q5189 entamoeba h
Q4dc4 trypanosoma
Q38Fn3 trypanosoma

ALIGNMENTS

Nakai K., Yada T., Nakamura Y., Ohara O., Isegai T., Sugano S.;
 RT "Complete sequencing and characterization of 21,243 full-length human
 RT cDNAs";
 RT Nat. Genet. 36:40-45(2004).
 [4] IDENTIFICATION BY MASS SPECTROMETRY, AND SUBCELLULAR LOCATION.
 MEDLINE=22317277; PubMed=12428849; DOI=10.1091/mbc.E02-05-0271;
 Scherl A., Coute Y., Deon C., Calle J.-A., Kindbeiter K., Sanchez J.-C.,
 Scherl A., Hochstrasser D.F., Diaz J.-J.;
 "Functional proteomic analysis of human nucleolus. ";
 Mol. Biol. Cell 13:4100-4109(2002).
 RRL -!- INTERACTION:
 086W57:XRN1; NBEXP1; Intact=EBI-373098, EBI-372406;
 CC -!- SUBCELLULAR LOCATION: Nuclear; nucleolar;
 CC -!- SIMILARITY: Contains 1 HEAT repeat.
 CC Copyrighted by the UniProt Consortium, see <http://www.uniprot.org/terms>
 CC Distributed under the Creative Commons Attribution-NoDerivs License
 CC
 DR EMBL; AX067150; CAC26776_1; ; Unassigned DNA.
 DR EMBL; AL359921; CAII3775_1; ; Genomic DNA.
 DR EMBL; AK001221; BAA91564_1; ALT_INIT; mRNA.
 DR IntAct; Q9P583; ;
 SWISS-2DPAGE; Q9H583; HUMAN.
 DR Ensembl; ENSG0000011925; Homo sapiens.
 DR HGNC; HGNC:2517; HEATR1.
 DR GO: GO:0005515; F:protein binding; IPI.
 DR InterPro; IPR012954; BP28_C.
 DR Intertierro; IPR000357; HEAT.
 DR Pfam; PF03146; BP28CT; 1.
 DR Pfam; PF02985; HEAT; 4.
 DR PROSITE; PS50077; HEAT_REPEAT; FALSE_NEG.
 DR KW Nuclear protein; Polymorphism.
 FT CHAIN 1 2144 HEAT repeat-containing protein 1.
 FT REPEAT 2106 2142 /FTid=PRO_00000186201.
 FT VARIANT 1694 1694 N -> S /FTid=VAR_010939.
 FT VARIANT 1854 1854 V -> A.
 FT VARIANT 1967 1967 /FTid=VAR_010940.
 FT VARIANT 2017 2017 N -> D /FTid=VAR_010941.
 FT VARIANT 348 348 E -> G, /FTid=VAR_010942.
 FT CONFLICT 367 607 H -> M /in Ref. 1).
 FT CONFLICT 607 607 M -> V (in Ref. 1).
 SQ SEQUENCE 2144 AA; 242370 MW; 7590BEA17A5FP39D CRC64;
 Query Match 99.9%; Score 8204; DB 1; length 2144;
 Best Local Similarity 99.9%; Pred. No: 0; Gaps 0;
 Matches 1627; Conservative 1; Mismatches 1; Indels 0; Gaps 0;
 QY 1 MTSIAQQLRLLAPQSDASLISRDVEASLFDPEAKATIDRATAIGCTGLEELKGIDP 60
 Db 1 MTSIAQQLRLLAPQSDASLISRDVEASLFDPEAKATIDRATAIGCTGLEELKGIDP 60
 QY 61 SREQQFEAPLFSLQAKTILERSVTKAVKQOLDENISLFLHSPFLPKAQCKLEWLHR 120
 Db 61 SREQQFEAPLFSLQAKTILERSVTKAVKQOLDENISLFLHSPFLPKAQCKLEWLHR 120
 QY 121 FHHLHYNODSLIACVLPHETRILFVRVIOLKINNSKRWRFLPPVKQSGVPLAKGTLIT 180
 Db 121 FHHLHYNODSLIACVLPHETRILFVRVIOLKINNSKRWRFLPPVKQSGVPLAKGTLIT 180
 QY 181 HCYKDGLGPMDFCISLVLTKSVKVFAYEGSSAQLRVLIAYFYSTIVSVALVAEDVDNTIA 240
 Db 181 HCYKDGLGPMDFCISLVLTKSVKVFAYEGSSAQLRVLIAYFYSTIVSVALVAEDVDNTIA 240
 .241 KLFPPYIQKGLKESLPLDRAATMICOISKVMTMIFVNSLASQIKLTQKPSLKDQ 300
 Db 241 KLFPPYIQKGLKESLPLDRAATMICOISKVMTMIFVNSLASQIKLTQKPSLKDQ 300

Run on: June 29, 2005, 00:46:09 ; Search time 28.4956 Seconds
 OM protein - protein search, using sw model

Total number of hits satisfying chosen parameters: 650591

Minimum DB seq length: 0 (without alignments)
 Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
 Maximum Match 100%
 Listing first 45 summaries

database : Issued Patents AA:*

1: /EMC_Celerra_SIDS3/pctdata/2/iaa/5_COMB.pep: *
 2: /EMC_Celerra_SIDS3/pctdata/2/iaa/_COMB.pep: *
 3: /EMC_Celerra_SIDS3/pctdata/2/iaa/H_COMB.pep: *
 4: /EMC_Celerra_SIDS3/pctdata/2/iaa/PCTUS_COMB.pep: *
 5: /EMC_Celerra_SIDS3/pctdata/2/iaa/RE_COMB.pep: *
 6: /EMC_Celerra_SIDS3/pctdata/2/iaa/backfiles.pep: *
 7: /EMC_Celerra_SIDS3/pctdata/2/iaa/backfiles.pep: *

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

result No.	Score	Query Match Length	DB ID	Description
1	310	3.8	200	US-09-248-796A-15108
2	228	2.8	3433	Sequence 15108, A
3	228	2.8	3433	Sequence 10, Appl
4	227.5	2.8	3433	Sequence 1136, Ap
5	227.5	2.8	2954	Sequence 1, Appl
6	226.5	2.7	3830	Sequence 1, Appl
7	216.5	2.6	3829	Sequence 16, Appl
8	215	2.6	2663	Sequence 12:2, Ap
9	214.5	2.6	3829	Sequence 2, Appli
10	212.5	2.6	3829	Sequence 31, Appl
11	212	2.6	2108	Sequence 87, Appl
12	206.5	2.5	1786	Sequence 8, Appli
13	206.5	2.5	2482	Sequence 6, Appli
14	206.5	2.5	3248	Sequence 1, Appli
15	206.5	2.5	3248	Sequence 1, Appli
16	203.5	2.5	3210	Sequence 154, Ap
17	202	2.5	3066	Sequence 12, Appli
18	201.5	2.5	1979	Sequence 6368, Ap
19	201.5	2.5	2047	Sequence 704, Ap
20	201.5	2.5	10182	Sequence 3159, Ap
21	198	2.4	3079	Sequence 4, Appli
22	197.5	2.4	2186	Sequence 10828, A
23	197.5	2.4	2349	Sequence 914, App
24	194.5	2.4	4872	Sequence 6368, Ap
25	194	2.4	2	US-09-949-016-7404
26	194	2.4	3056	Sequence 10280, A
27	194	2.4	3056	Sequence 8, Appli

RESULT 1
 US-09-248-796A-15108
 Sequence 15108, Application US/09248796A
 Patent No. 6747137

GENERAL INFORMATION:

APPLICANT: Keith Weinstock et al

TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO CANDIDA ALBICIAN

FILE REFERENCE: 107196.132

CURRENT APPLICATION NUMBER: US/09/248,796A

CURRENT FILING DATE: 1999-02-12

PRIOR APPLICATION NUMBER: US 60/074,725

PRIOR FILING DATE: 1998-02-13

PRIOR APPLICATION NUMBER: US 60/096,409

PRIOR FILING DATE: 1998-08-13

NUMBER OF SEQ ID NOS: 28208

SEQ ID NO: 15108

LENGTH: 200

TYPE: PRT

ORGANISM: Candida albicans

; US-09-248-796A-15108

Query Match 3.8%; Score 310; DB 2; Length 200;
 Best Local Similarity 38.0%; Pred. No. 7.8e-18; Mismatches 70; Indels 18; Gaps 5;
 Matches 76; Conservative 36; Mismatches 18; Gaps 5;

Qy , 1 MTSLAQQLR-----ALPQDASLILSRDEVASLLEPDPEKATIDRTAAGCTGL 52
 Db , 4 MSLSAQQLQINKEKTAVASVQLDRKGSKLHSR----SLLFDPKQATOQDYEYIATEGL 59

Qy , 53 BELIGIDPSFPEQQFAPLISQAKTLERSVQTKAVNQKLDENISFLIHLSPYFLPKAQK 112
 Db , 60 EDLDELDSPRFNKFLKTSETSVNLDRNQTKDVIQLDKNIDAFLTVGPGYGLTSLK 119

Qy , 113 CLEMLIHRHILHNODSIACTACVLPYHETRIFPVRLQKLNNSKHRWFLWLPYK-OSGV 171

Db , 120 AVEMLVRRFHANKELMILTAFLPYFOPVFWLNVPKQNPQIFEWLGVYKDLKT 179

Qy , 172 PLAKGLITHCYKDGFMDP 191

Db , 180 PPASSIL----KGMSAMNP 194

RESULT 2
 US-09-911-501B-10
 Sequence 10, Application US/09091501B
 Patent No. 6515413

GENERAL INFORMATION:

APPLICANT: Tibley, Jonathon M

APPLICANT: Davies, Kay E
 TITLE OF INVENTION: Utrophin gene expression
 FILE REFERENCE: 620-42
 CURRENT APPLICATION NUMBER: US/09/091,501B
 CURRENT FILING DATE: 1998-06-18
 PRIOR APPLICATION NUMBER: PCT/GB96/03156
 PRIOR FILING DATE: 1996-12-19
 PRIOR APPLICATION NUMBER: GB 9525962.8
 PRIOR FILING DATE: 1995-12-19
 PRIOR APPLICATION NUMBER: GB 9615197.9
 PRIOR FILING DATE: 1996-07-26
 PRIORITY APPLICATION NUMBER: GB 9622174.2
 NUMBER OF SEQ ID NOS: 15 2.1
 SEQ ID NO: 10
 LENGTH: 3433
 TYPE: PRT
 ORGANISM: Artificial Sequence
 FEATURE:
 NAME/KEY: misc feature
 LOCATION: (239) .. (250)
 OTHER INFORMATION: Description of Artificial Sequence: Full length
 OTHER INFORMATION: utrophin construct; Xaa = unknown
 US-09-091-501B-10

Query Match Best Local Similarity 2.8%; Score 228; DB 2; Length 3433;
 Matches 314; Conservative 272; Mismatches 565; Indels 528; Gaps 78;

QY	144 FVRVIIQLKINNSHRWFMLLPVKQS--GVPLAKGTJLTHCYRDGLGMDIFCSSLVTKSV 200
Db	804 FKQDDELEVKIKTEKEWVKHTSISSESSRSQSLPSIKDS---CORBL----TNLLGHP 853
Qy	201 KVFFEYGPSSAQRLVLLAYASTTVAALSAVADSDNTIAKLFPEYIQLKGKSSPDYKA 260
Db	854 KI-----EMARACCSALMSQPSAD-----FVQRGFDSPFLGRYAV 889
Qy	261 TYMILCQISV--KVMENTFVNLSAQIKLTKEPLKISLKDJSCLIVLQRQPKESLG 317
Db	890 -----QEAVEDRQHOLENLKGPGHAYLETUICLVDLNS-----ENKAQVSLN 935
Qy	318 KKPFPHLCNVPDLITLHGSETDVSPYLRYMLPHLVVSIIHVTGEETEGMD---- 371
Db	936 V-----LNDLKVEKALQEKTIDELLENQKPAL----HKAEBTAKALEKNVHD 981
Qy	372 -GOIYKRHLAELTKISKLKNNDLILLASLFEETISYSSEEMNSNKYLINQFLPJR 430
Db	982 VEKLYKQEFDDVQCK-----WNKLKVVLVSKDHILLEELALT-LR 1019
Qy	431 LLESKYPRILDVVBEHLIKEAD-LKKQFLFHQVSLSTSGGKVQFLADSPTSLMSLNH 489
Db	1020 AFEAD----STVIEKWMGDVGVKDFLMKO-----QAQCDDAGSLQRQOLDO 1058
Qy	490 PLADPVRL---AMNLKKIMKTSKEGIDESFTKEAVIARLGDDNIDVLSAISAFEP 544
Db	1059 CSATVNEIETIESLKSOMKEIETNLURSG-PVAGIKTWQTRIGD-----YDQ 1105
Qy	545 KEHHSSEVTI--SNLNLFORA----ELSNGHEN----- 573
Db	1106 LEKUSKEIATQOKSRSLSESOKANLKKDIAEMQIMTOABEEYERDFYKSPRELEAV 1165
Qy	574 --EVKIAADILIKE---ETLSEN-----DQSNQVVCULLPFWVNN----- 611
Db	1166 EEMRKRAKEDVLQKEVVKLIKDNKLAAKVPSP3QELTSLENVQOLLCKRIRGKC 1225
Qy	612 --DTESEAEMKIAIYLSKGICSHPLRGWEALENTIKSTK--PGKLIGV----AN 651.
Db	1226 HTLBEVWSWCIELHLYD-----LETTWNTLEERMSTVLPKEKTDVANEASLS 1276
Qy	662 QRMIBELLADN---INLGDPSSMLKVEDLISVBESSENLVQKUTFHILVIVSC--- 714
Db	1277 ESVLRHPADNRTOIRELGOTLIDGILDDPIIS-KELEAFNSRYBDLHLAESQISLEQ 1135

QY 715 CSSLKEITHPPFAIRVFSLIQLKKKLGESVT-----AVIPSEWHIELDRGIVPE 766
 Db 1336 LQVIRETD---OMLQVLOBSQELGDQKQTLTYLDRDAFQVPOE-----AOQIAOE 1383
 QY 767 LWAH--YVEELNSTOR---VAVEDSVLFLVFSLKRFTALKAPRSFPGKDIAMNPQEQLD 821
 Db 1384 ISAHETLLEELRRNMRSQPLTSPESRTARGSSQMDVQKRLREVSTIKEFQLOFKPANEQR 1443
 QY 822 SRYLHLIGL-FEMMINGADAVHFRYTMKLFKVLHEDVOLFKEFCSVLNTYGSLSNP 880
 Db 1444 MLDCKRVLDGVKAELHVLDVKDVPD-----T0THLJKCMVLYKTS-----E 1487
 QY 881 LNC SVKTVLQTOALIVGCAMLSQKTO---CKHQ LASISPSVTSU LINGSPVKE-- 933
 Db 1488 VKL EYETVIK-----GRHIVGKQQTDPNPKGMDQBLTSIK----VLYNDIGAQVTBGQ 1537
 Db 934 -----VRAAIQCLQALSGVASPFYIILHISK--AEITSDAVNIQIATL 980
 Db 1538 DLERASQALARQKMGKEAASLSEWISAT-----EVQKSTSEGJGDLPEISWAKN 1590
 Qy 981 FEBOLOREK-KLKSHQKASETTLKNU-----CV----YSCRPSYIAKOLMKVIGV 1025
 Db 1591 LKDLERKKAQNNTYESAALQNLIESSEPLBERLCLVNLAGNSRVTWTEDWCNTLASH 1650
 Qy 1026 NGEMV-----LSQLLPMASOLLEKQEKPTAVLKDTEAMYLHLTGKYNPSVSLINE 1077
 Db 1651 QNQLETFDGNVAMHISTWLYOAEALLDEKEPTS---KOBETVRL-VSELDDANLOVEN 1707
 Qy 1078 DPKSLDITKAVHNTKELYASHTQTIQTALEKIKTPPFAASIDEKVQO---KLU---R 1129
 Db 1708 RDOALIUMNARASSSSRBLVE-----PKLABLNRF-----EKVSQHKSACKLAQE 1754
 Qy 1130 MFLDLVNCNSKHCQTVSSVFKGISVNAEQRILEPDPDKAKPLGTWQKQRQKOMQKK 1189
 Db 1755 PIQCLVTTETFGVPSDD-----EKLENDIENMLK-----FVEKHLESBDE 1800
 Qy 1190 SQDLESVOEFGGSYWQRTVLTBLLH-----KKKURSPQIILVPTLFLNLSR 1236
 Db 1801 KMDBEESAO---TIEFLVORGEEMHJQPMEDNKKEKIRLQLLILHTRYNK-- 1845
 Qy 1237 CLEPLPQROGQHMYTKQLLQKLSCLNICKLSPDGKPKDILDEEKNVNLIVQTCULSE 1296
 Db 1846 -IKAPIQORKONGQLASGIRSSL-----PTDYLVE---INKLICMDDVE 1887
 Qy 1297 MPQTHHALLLGLTVAGIFPPDKVLUHNMSIFTFMGAMVRLD---DTYSQVINKVMM 1352
 Db 1888 -----LSLNUPELMATAIYDFSFQ----- 1906
 Qy 1353 VIPALIQOSDSGSIEVSRNVEETVVKI---ISVFDVDALPHVPEHRRPLPILVQVUDTJGAE 1409
 Db 1907 -----EDSL---KKNKOOLDKULGEQOAVIHEROPDVILEASGEPAIQRDTJQL 1953
 Qy 1410 KFLWILLIPLFEGQYVTKVLLAYAYGKDAILBADERWFSCCEFSVQHOIOSMLNIIQY 1469
 Db 1954 NAKW-----DRINRMTSDRKGCFDRAMEEWRFQHCD-----LNDLTQW 1991
 Qy 1470 LKLPEEKEPTPK--AVSFNKSSESQEMLOVFNVEHTSKQRLRHFPLSVSPMSQLISS 1527
 Db 1992 ITBEEELLVDTCPFGSSIDUEKARIHQBLVE-GISS-----OPSTAALNITG 2039
 Qy 1528 NNFLKVKVVESSG--PELKGLBERLLETVLYGIVSANOQMER---NADLTVKFRA 1579
 Db 2040 DGIVQKUSQADSFLKEKLGAJOR-----WDAIVAEVKORQPRLKQGESKOVKMYRHO 2092
 Qy 1580 L-----LISKAYLUDKVNALPPTETEPVIRGLV-----GNPLPSYRKALDLINK 1626
 Db 2093 IDEICWITKAHAMQKESTTELGEN-LQELRDLTQEMEVHAELKWMNRTELEMSDLK 2150
 RESULT 3
 US-09-538-092-1136
 ; Sequence 1136, Application US/09538092

GenCore version 5.1.9
 Copyright (c) 1993 - 2006 Biocceleration Ltd.

Om protein - protein search, using sw model
 Run on: June 29, 2006, 00:52:34 ; Search time 101.893 Seconds
 (without alignments)

Perfect score: 7405.549 Million cell updates/sec
 Sequence: 1 MISLAQQQLQRALAPQSDASL.....NPLPSVRRKALDILANKLQQ 1629

Title: US-09-603-665-5_COPY_1_1629
 Perfect score: 8212

Scoring table: BIOSUM62
 Gapop 10.0 , Gapext 0.5
 Searched: 2097797 seqs, 463214858 residues

Total number of hits satisfying chosen parameters:
 Minimum DB seq length: 0
 Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
 Maximum Match 100%
 Listing first 45 summaries

Database : Published Applications AA Main: *
 1: /EMC_Celerra_SIDS3/prodata/2/pupbaa/US10 PUBCOMB.pep: *
 2: /EMC_Celerra_SIDS3/prodata/2/pupbaa/US08 PUBCOMB.pep: *
 3: /EMC_Celerra_SIDS3/prodata/2/pupbaa/US07 PUBCOMB.pep: *
 4: /EMC_Celerra_SIDS3/prodata/2/pupbaa/US10A PUBCOMB.pep: *
 5: /EMC_Celerra_SIDS3/prodata/2/pupbaa/US11 PUBCOMB.pep: *
 6: /EMC_Celerra_SIDS3/prodata/2/pupbaa/US11_PUBCOMB.pep: *

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match Length	DB ID	Description	
1	8212	10.0	2144	5 US-10-445-237-230	
2	819.8	2144	6 US-11-124-368A-277	Sequence 230, App	
3	7635	93.0	2044	6 US-11-124-368A-278	Sequence 277, App
4	7627.5	92.9	1569	5 US-10-450-763-45591	Sequence 45591, A
5	7255	88.5	2036	6 US-11-124-368A-280	Sequence 276, App
6	7266	88.5	2036	6 US-11-124-368A-281	Sequence 280, App
7	7266	88.5	2036	6 US-11-124-368A-281	Sequence 281, App
8	3694	45.0	2734	4 US-10-128-714-3300	Sequence 330, App
9	2824.5	34.4	1149	5 US-10-128-558-167	Sequence 167, App
10	1916	23.3	897	4 US-10-764-422-177	Sequence 177, App
11	1603	19.5	325	5 US-10-128-559-358	Sequence 358, App
12	1396.5	19.5	2096	5 US-10-745-237-56	Sequence 36, Appl
13	1396.5	17.0	1709	6 US-11-097-143-22518	Sequence 22518, A
14	1095.6	8.6	1798	4 US-10-128-714-3388	Sequence 8388, App
15	704.5	8.6	1814	4 US-10-128-714-3388	Sequence 3388, App
16	684.5	8.3	1818	4 US-10-032-583-7616	Sequence 7616, App
17	674.5	8.2	2122	4 US-10-437-963-189782	Sequence 189782, App
18	663	8.1	135	4 US-10-264-049-2412	Sequence 2411, App
19	379	4.6	77	3 US-09-864-761-37036	Sequence 37036, A
20	237	2.9	952	4 US-10-424-593-248111	Sequence 248111, App
21	228	2.8	3433	4 US-10-408-768A-731	Sequence 731, App
22	228	2.8	3433	4 US-10-408-765A-732	Sequence 732, App
23	228	2.8	3433	5 US-10-766-149-5681	Sequence 5681, App
24	228	2.8	4131	4 US-10-369-493-5136	Sequence 5136, App
25	227.5	2.8	2954	5 US-10-650-280-1	Sequence 1, Appl
26	225.5	2.7	1583	4 US-10-408-768A-1635	Sequence 1635, App
27	218.5	2.7	2543	5 US-10-828-985A-9	Sequence 9, Appl

ALIGNMENTS

RESULT 1
 US-10-745-237-230
 Sequence 230 Application US10745237
 Publication No. US20050227301A1
 GENERAL INFORMATION:
 APPLICANT: Science Limited
 APPLICANT: Glover, David
 APPLICANT: Bell, Graham
 APPLICANT: Frenz, Lisa
 APPLICANT: Midgley, Carol
 TITLE OF INVENTION: Cell Cycle Progression Proteins
 FILE REFERENCE: P015819WO CYK
 CURRENT APPLICATION NUMBER: US 10/745,237
 CURRENT FILING DATE: 2003-12-23
 PRIOR APPLICATION NUMBER: US 60/439,123
 PRIOR FILING DATE: 2003-01-10
 PRIOR FILING DATE: 2003-05-06
 NUMBER OF SEQ ID NOS: 600
 SOFTWARE: Patentin version 3.1
 SEQ ID NO 230
 LENGTH: 2144
 TYPE: PRT
 ORGANISM: Homo sapiens
 OTHER INFORMATION: Q9H583
 US-10-745-237-230

Query Match Score 100 %; DB 5; Length 2144;
 Best Local Similarity 100 %; Pred. No. 0;
 Matches 1629; Conservative 0; Mismatches 0;
 Indels 0; Gaps 0;

US-10-745-237-230

QY 1 MISLAQQQLQRALAPQSDASLISRDEVASLIEDPKREATIDDTAFAGTCLEELIGIDP 60
 Db 1 MISLAQQQLQRALAPQSDASLISRDEVASLIEDPKREATIDDTAFAGTCLEELIGIDP 60

QY 61 SIEQFEAPLFSLQAKTLERSQTAKNQOLDENISFLIHSIPLYFLIKPAOKCLELIR 120
 Db 61 SIEQFEAPLFSLQAKTLERSQTAKNQOLDENISFLIHSIPLYFLIKPAOKCLELIR 120

QY 121 FHILHNQDSIACVLPYHETRIFVVIOLKINNSKRHRWILLPLKOSGFLAKTIT 180
 Db 121 FHILHNQDSIACVLPYHETRIFVVIOLKINNSKRHRWILLPLKOSGFLAKTIT 180

QY 181 HCKDQLGPMDFICSLSVTKSVYFAEPGSSQLRVLVAFYSTIVSVALVAEDVSNIA 240
 Db 181 HCKDQLGPMDFICSLSVTKSVYFAEPGSSQLRVLVAFYSTIVSVALVAEDVSNIA 240

QY 241 KLLPYIQLGLKLSPDQYRAATYMICQISVYMENTFVNLSIASQIKLTKPLSIKD 300
 Db 241 KLLPYIQLGLKLSPDQYRAATYMICQISVYMENTFVNLSIASQIKLTKPLSIKD 300

Db 241 KLFPIYIQGKSLPDIRATMICQISVKVMENTFVNSLASQIQLIKTIPSLIKD 300
 Qy 301 LSCIVLILQRQKPSLGKPFPHCNVPDLITHGISETYDPSPLRMLPHWSTH 360
 Db 301 LSCIVLILQRQKPSLGKPFPHCNVPDLITHGISETYDPSPLRMLPHWSTH 360
 Qy 361 HVTGEGETEGMDGOYKRHLAILTAKISKNOLNLDHLASLFEELYISYSOEEMDSNKYL 420
 Db 361 HVTGEGETEGMDGOYKRHLAILTAKISKNOLNLDHLASLFEELYISYSOEEMDSNKYL 420
 Qy 421 LNEFLPLRLLESKIPRTLDVWEELHLKIAJDKQELFHOFVSLSTSGGKQFLAND 480
 Db 421 LNEFLPLRLLESKIPRTLDVWEELHLKIAJDKQELFHOFVSLSTSGGKQFLAND 480
 Qy 481 TSLMLSLNLHPLAPYRILAMNHKLKIMTKTSKEGVRSFIKAVALRGDDNDVVLSA 540
 Db 481 TSLMLSLNLHPLAPYRILAMNHKLKIMTKTSKEGVRSFIKAVALRGDDNDVVLSA 540
 Qy 541 FEIKKEHESSEVT-SNLNLFORALBLSKONGEWYEVKIADILKEEILSENDOLSNQV 600
 Db 541 FEIKKEHESSEVT-SNLNLFORALBLSKONGEWYEVKIADILKEEILSENDOLSNQV 600
 Qy 601 VCLIPFWINDDTBESAEMKAIATLUSKGICSLHPLRGWEAEALENVKSTPKGLIGA 660
 Db 601 VCLIPFWINDDTBESAEMKAIATLUSKGICSLHPLRGWEAEALENVKSTPKGLIGA 660
 Qy 661 NORMIELLADNINGGDPSMILKOMTDLISGEERSFNLKOKVTPHVISVLUCCSIKE 720
 Db 661 NORMIELLADNINGGDPSMILKOMTDLISGEERSFNLKOKVTPHVISVLUCCSIKE 720
 Qy 721 THFFPAIRVSLSLKKIKKLESVTAVIPSEWHIELMDRGIPVELWAHYVEBLNSTO 780
 Db 721 THFFPAIRVSLSLKKIKKLESVTAVIPSEWHIELMDRGIPVELWAHYVEBLNSTO 780
 Qy 781 VAYEDSVLVEFSLKFKYIYAKAPSFKPKDIDWWPEQLEDSRYSYLHLIGLEMMINGA 840
 Db 781 VAYEDSVLVEFSLKFKYIYAKAPSFKPKDIDWWPEQLEDSRYSYLHLIGLEMMINGA 840
 Qy 841 DAVFPRVLMKLFKTHLEVFQKFCSVLWTVSSLSLPLNCSSVKTQTLQIYGCAM 900
 Db 841 DAVFPRVLMKLFKTHLEVFQKFCSVLWTVSSLSLPLNCSSVKTQTLQIYGCAM 900
 Qy 901 LSSQTKTOCKHOLASSISSPVTSLSLINLGSPVKEYRAAOCLQALSGVASPFVLLIDH 960
 Db 901 LSSQTKTOCKHOLASSISSPVTSLSLINLGSPVKEYRAAOCLQALSGVASPFVLLIDH 960
 Qy 961 SKABEITSDAAYVODLATLFEEREKLUKSHOKLSETTLKNUCVCSPSKYAKDMK 1020
 Db 961 SKABEITSDAAYVODLATLFEEREKLUKSHOKLSETTLKNUCVCSPSKYAKDMK 1020
 Qy 1021 VLOCYNGEMVLSQLPMAQOLLEKIQEPTAVLQDEAMWLHTLGKMFHFSVLNEPK 1080
 Db 1021 VLOCYNGEMVLSQLPMAQOLLEKIQEPTAVLQDEAMWLHTLGKMFHFSVLNEPK 1080
 Qy 1081 SLDTFIKAVHTKELYAGMPTQITALEKITKPPFAALSDEKVOKLRLFDLUNCKN 1140
 Db 1081 SLDTFIKAVHTKELYAGMPTQITALEKITKPPFAALSDEKVOKLRLFDLUNCKN 1140
 Qy 1141 SHCQTVSSVFKGSVNAQVRILEPPDKAKPPTVQOKRQHQKQSDSLBSVQEG 1200
 Db 1141 SHCQTVSSVFKGSVNAQVRILEPPDKAKPPTVQOKRQHQKQSDSLBSVQEG 1200
 Qy 1201 GSYMORVTLLELLOHKKULSRPOLLVPLTFLNLSRCLEPLPQOGNNMEXTQDLSCL 1260
 Db 1201 GSYMORVTLLELLOHKKULSRPOLLVPLTFLNLSRCLEPLPQOGNNMEXTQDLSCL 1260
 Qy 1261 NICOKLSPGKIKDILBEKFVELVOCIRSSEMPOTHHALLGTVAGFPDKV 1320
 Db 1261 NICOKLSPGKIKDILBEKFVELVOCIRSSEMPOTHHALLGTVAGFPDKV 1320
 Qy 1321 HNTINSIFTMGANTMRDDTYSFQVINKTVKMFVPAIQLQSDSGSIE/SRNVERBVKII 1380
 Db 1321 HNTINSIFTMGANTMRDDTYSFQVINKTVKMFVPAIQLQSDSGSIE/SRNVERBVKII 1380

RESULT 2
 US-11-124-368A-277

; Sequence 277, Application US/11124368A
 ; Publication No. US20050287559A1

; GENERAL INFORMATION:

; APPLICANT: Michele Carrill

; APPLICANT: James J Devlin

; APPLICANT: May Luke

; TITLE OF INVENTION: Genetic Polymorphisms Associated with

; Vascular Diseases, Methods of Detection and Uses Thereof

; FILE REFERENCE: CLO1524

; CURRENT APPLICATION NUMBER: US/11/124, 368A

; CURRENT FILING DATE: 2005-05-09

; PRIORITY NUMBER: US 60/568, 845

; PRIORITY FILING DATE: 2004-05-07

; PRIORITY APPLICATION NUMBER: US 60/625, 936

; PRIOR FILING DATE: 2004-11-09

; NUMBER OF SEQ ID NOS: 2112

; SOFTWARE: FastSEQ for Windows Version 4.0

; SEQ ID NO: 277

; LENGTH: 2144

; TYPE: PRT

; ORGANISM: Homo sapiens

US-11-124-368A-277

Query Match 99.8%; Score 8195; DB 6; Length 2144;

Best Local Similarity 99.8%; Pred. No. 0; Indels 0; Gaps 0;
 Matches 1626; Conservative 1; Mismatches 2;

Db 1 MTSLAQIQLRQLAPOSQASLRSDEVASLFLPKEATIDRUTAFATCTGIGLELLIDP 60
 Db 1 MTSLAQIQLRQLAPOSQASLRSDEVASLFLPKEATIDRUTAFATCTGIGLELLIDP 60
 Qy 61 SFROPEAPLFSQALKLTERSVOVKANQOLDENISFLIHSPYFLIKPAQKCLEWLR 120
 Db 61 SFROPEAPLFSQALKLTERSVOVKANQOLDENISFLIHSPYFLIKPAQKCLEWLR 120
 Qy 121 FHLYNODSLIACVLPHTETLTVYRVIOLKLNNSKRWRWFLIPVKOSGVPLAKGTIT 180
 Db 121 FHLYNODSLIACVLPHTETLTVYRVIOLKLNNSKRWRWFLIPVKOSGVPLAKGTIT 180
 Qy 181 HCYKDGLGMDPCLSLVTKSYKVAEYCGSSAOLRVLAFYASTIVSALVAQSDNTIA 240
 Db 181 HCYKDGLGMDPCLSLVTKSYKVAEYCGSSAOLRVLAFYASTIVSALVAQSDNTIA 240
 Qy 181 HCYKDGLGMDPCLSLVTKSYKVAEYCGSSAOLRVLAFYASTIVSALVAQSDNTIA 240
 Db 241 KLFPIYIQGKSLPDIRATMICQISVKVMENTFVNSLASQIQLTIPSLIKD 300
 Qy 241 KLFPIYIQGKSLPDIRATMICQISVKVMENTFVNSLASQIQLTIPSLIKD 300
 Db 241 KLFPIYIQGKSLPDIRATMICQISVKVMENTFVNSLASQIQLTIPSLIKD 300
 Qy 301 LSCIVLILQRQKPSLGKPFPHCNVPDLITHGISETYDPSPLRMLPHWSTH 360

Copyright (c) 1993 - 2006 Biocceleration Ltd.

GenCore version 5.1.9

OM protein - protein search, using sw model

Title: US-09-603-665-5_COPY_1_1629
Perfect score: 8212
Sequence: 1 MTSLAQQLQRALPOSDSL... NPLPSVRKALLNNKLOQ 1629

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 101426 seqs, 23619683 residues

Total number of hits satisfying chosen parameters: 103426

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Published Applications AA New: *

1: /EMC_Celerra_SD33/ptodata/2/pubpaas/US09_NEW_PUB_pep: *
2: /EMC_Celerra_SD33/ptodata/2/pubpaas/US07_NEW_PUB_pep: *
3: /EMC_Celerra_SD33/ptodata/2/pubpaas/US08_NEW_PUB_pep: *
4: /EMC_Celerra_SD33/ptodata/2/pubpaas/PCT_NEW_PUB_pep: *
5: /EMC_Celerra_SD33/ptodata/2/pubpaas/US10_NEW_PUB_pep: *
6: /EMC_Celerra_SD33/ptodata/2/pubpaas/US11_NEW_PUB_pep: *
7: /EMC_Celerra_SD33/ptodata/2/pubpaas/US60_NEW_PUB_pep: *
8: /EMC_Celerra_SD33/ptodata/2/pubpaas/US60_NEW_PUB_pep: *

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Match	Query Length	DB ID	Description
1	3694	45.0	734	7	US-11-293-697-3300
2	206.5	2.5	3113	6	US-10-505-928-325
3	188.5	2.3	2209	7	US-11-301-554-1903
4	171	2.1	1325	6	US-10-449-902-51794
5	166.5	2.0	1206	6	US-10-449-902-43305
6	165	2.0	1176	6	US-10-449-902-46322
7	163.5	2.0	1328	6	US-10-504-973-32
8	159.5	1.9	2671	6	US-10-505-928-784
9	158.5	1.9	1116	6	US-10-449-902-47106
10	157.5	1.9	980	7	US-11-283-692-296
11	157	1.9	1205	7	US-11-293-692-2967
12	155	1.9	764	6	US-10-449-902-45961
13	152.5	1.9	1230	7	US-11-283-102-378
14	151.5	1.8	1218	6	US-10-449-902-47145
15	150	1.8	1020	6	US-10-471-571A-4624
16	147.5	1.8	1443	6	US-10-486-020-14
17	146.5	1.8	116	6	US-10-471-571A-4478
18	146.5	1.8	1056	6	US-10-511-937-2929
19	141.5	1.7	1345	6	US-0-449-902-55021
20	141	1.7	1345	6	US-0-449-902-55021
21	140.5	1.7	1663	6	US-10-505-928-227
22	139.5	1.7	1174	6	US-10-449-902-9996
23	139.5	1.7	1345	6	US-10-449-902-55428
24	139.5	1.7	2364	7	US-11-289-102-242
25	136.5	1.7	1575	6	US-10-505-928-257

ALIGNMENTS

RESULT 1
US-11-293-697-3300
; Sequence 3300, Application US/11293697
; Publication No. US20060105376A1
; GENERAL INFORMATION:
; APPLICANT: HELIX RESEARCH INSTITUTE
; TITLE OF INVENTION: Novel full length cDNA
; FILE REFERENCE: H1-A0106
; CURRENT APPLICATION NUMBER: US/11/293, 697
; CURRENT FILING DATE: 2005-12-05
; PRIOR APPLICATION NUMBER: US/10/108, 260
; PRIORITY FILING DATE: 2002-03-28
; NUMBER OF SEQ ID NOS: 5458
; SOFTWARE: PatentIn Ver. 2.1
SEQ ID NO 3300
LENGTH: 734
TYPE: PRT
ORGANISM: Homo sapiens
US-11-293-697-3300

Query Match 45.0%; Score 3694; DB 7; Length 734;

Best Local Similarity 100.0%; Pred. No. 1.7e-208;
Matches 734; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Sequence 141, App
Sequence 45252, A
Sequence 2392, AP
Sequence 45594, A
Sequence 37528, A
Sequence 3282, APP
Sequence 343, APP
Sequence 494, APP
Sequence 43318, A
Sequence 43257, A
Sequence 36405, A
Sequence 154, APP
Sequence 3910, APP
Sequence 2968, APP
Sequence 2, APP
Sequence 351, APP
Sequence 155, APP
Sequence 7, APP
Sequence 102, APP
Sequence 140, APP
Sequence 142, APP

Sequence 141, App
Sequence 45252, A
Sequence 2392, AP
Sequence 45594, A
Sequence 37528, A
Sequence 3282, APP
Sequence 343, APP
Sequence 494, APP
Sequence 43318, A
Sequence 43257, A
Sequence 36405, A
Sequence 154, APP
Sequence 3910, APP
Sequence 2968, APP
Sequence 2, APP
Sequence 351, APP
Sequence 155, APP
Sequence 7, APP
Sequence 102, APP
Sequence 140, APP
Sequence 142, APP

QY 361 HVTGEBETEGMDQIYKRRHLEALITKISLKNNDLHDJHLASLAFEEYSYSSOBEDMSKVSL 420
 361 HVTGEBETEGMDQIYKRRHLEALITKISLKNNDLHDJHLASLAFEEYSYSSOBEDMSKVSL 420
 Db 421 LNEOPLPLTRLESKPYRTDVLVUERHLKEJADLKHQELPHOFVULSTSGKCYOLADSD 480
 Qy 421 LNEOPLPLTRLESKPYRTDVLVUERHLKEJADLKHQELPHOFVULSTSGKCYOLADSD 480
 Db 421 LNEQFLPLTRLESKPYRTDVLVUERHLKEJADLKHQELPHOFVULSTSGKCYOLADSD 480
 Qy 481 TSLMSLNHPLAPVRLAMPHLKMKTSKEGVDFSIKEVALASLGDDMDVVAISA 540
 Db 481 TSLMSLNHPLAPVRLAMPHLKMKTSKEGVDFSIKEVALASLGDDMDVVAISA 540
 Qy 481 TSLMSLNHPLAPVRLAMPHLKMKTSKEGVDFSIKEVALASLGDDMDVVAISA 540
 Qy 541 FEIPEKHFSSEVTISNLINQFORAELSKNGEWEYUJKIAADLILKEBILSENDQSNQV 600
 Db 541 FEIPEKHFSSEVTISNLINQFORAELSKNGEWEYUJKIAADLILKEBILSENDQSNQV 600
 Qy 541 FEIPEKHFSSEVTISNLINQFORAELSKNGEWEYUJKIAADLILKEBILSENDQSNQV 600
 Db 601 VCLLPPVVIINDDTSAEMKIAYISKSGICSLHPLURGEALENVKSTKPGKLIGVA 660
 Qy 601 VCLLPPVVIINDDTSAEMKIAYISKSGICSLHPLURGEALENVKSTKPGKLIGVA 660
 Db 601 VCLLPPVVIINDDTSAEMKIAYISKSGICSLHPLURGEALENVKSTKPGKLIGVA 660
 Qy 661 NORMMELLADNNINTGDPSSMLKVMVDLISVGEEESPNLKOKVTFHVLISVLVSSSLKE 720
 Db 661 NORMMELLADNNINTGDPSSMLKVMVDLISVGEEESPNLKOKVTFHVLISVLVSSSLKE 720
 Qy 721 THFPAIRVFSLQ 734
 Db 721 THFPAIRVFSLQ 734
 Qy 721 THFPAIRVFSLQ 734
 Db 721 THFPAIRVFSLQ 734

RESULT 2

US-10-505-928-325

Sequence 325, Application US/10505928

Publication No. US20060088532A1

GENERAL INFORMATION:

APPLICANT: Ludwig Institute for Cancer Research et al.

TITLE OF INVENTION: LYMPHATIC ENDOTHELIAL GENES

FILE REFERENCE: 28967/39178

CURRENT APPLICATION NUMBER: US/10/505, 928

PRIOR APPLICATION NUMBER: US 60/363, 019

PRIOR FILING DATE: 2002-03-07

NUMBER OF SEQ ID NOS: 866

SOFTWARE: PatentIn 3.2

SEQ ID NO 325

LENGTH: 3113

TYPE: PRT

ORGANISM: Homo sapiens

US-10-505-928-325

Query Match 2.5%; Score 206; DB 6; Length 3113; Best Local Similarity 18.4%; Pred. No. 0.0016; Matches 309; Conservative 294; Mismatches 584; Indels 489; Gaps 75;

Qy 266 CQISKVMTENTFNSLASQQLIKITKIPSLKDCISLCLVLLQPKPSIGKKPFHLC 325
 Db 1425 CQMSKMSLQTYVDSKAENLVSTNLNFGD---LVKEMOGLLEGI----- 1471.
 Db 326 NVPDLITLHGISETYDVSPLLRYMLPHLUVSIHHVTGEBETEGMDQIYKRRHLEALITK 385
 Db 1472 -----VPSLSSCVDSSSISLG-DSSYRALLQ-TGD 154

Qy 386 ISLKNNL-----DHLLASLIFEEYTSS----- 409
 Db 1505 MSLLSNLEGAVSANOCVSDVFCCSLSQEEENLTKTPSAPAKGVHELESCEVFRQSLK 1564
 Db 410 -QEENDSNKSYSLNNOFLPFLRLKSKPYRTDVLVUERHLKEJADLKHQELPHOFVULST 468
 Qy 1565 LEBRQESOGI-MKNGKTOELEOLUSE-ROEBCDCKRQYLS-E-----NEOWQOKLISVT 1616
 Db 469 SGGKXQFLADSDTSMLSLNHPLAPVRLAMPHLKMKTSKEGVDFSIKEVALASLGDDMDVVAISA 528
 Db 1617 LEMBKLAEBKKQTBLSLBLEVARLQ-----QGLDSL-----SRSLLGI 1657

Qy 529 DNDIVVLSAISAFETFKEHFSSEVT-----ISNLNLFORAESLKGNGEWEYUJKIAAD 581
 Db 1658 DTDEAIQGRNESCDISKEH-TSETTERTPKHDVHQICDKDQAQDIN---LDIKITET 1711
 Qy 582 ILIK-EISESENDOLSNQVWCLLPFWVVIINDDTSEASBNKIAATYLSKSGICSLPL-LR 638
 Db 421 LNEQFLPLTRLESKPYRTDVLVUERHLKEJADLKHQELPHOFVULSTSGKCYOLADSD 480
 Qy 639 GWEAEALENV--IKSPKGK-----IGVANQKMEILLAD----NINGDPSMLK- 683
 Db 1763 GNQEDITHNQLQRVKETSNENLRLHVIEDDRKVESLNLNEKELDSKLHQBQVQNTKIE 1822
 Db 1712 GALKPTGBCSGEQSDPTN-----IYPPEBDKTQGSSCIS-BLSFSGPNALPMDFL 1762
 Qy 684 ---VEDLISGEEESPNLKKQKNTF--HVLISVLVSCCSLK--ETHFFPAIRV-- 729
 Db 1823 ACIEBKINGBLKKENSIDSLEKEXFDCHQELHORVETSEGLNDSLEMADKSSRDIG 1882
 Qy 730 -----FSLQLOKKIKKLESVITAWEIPSEWHIIMLDRGIPVWLAWHVEELN 776
 Db 1883 DNVAKNDNSWERFLDVENELSRIRSEKASIE-----HEALYLEAD-----EVV 1927
 Qy 777 STORYAV---EDSVPLVFSLKKFTYALKAKPSFKRGDI----WWNPQROKEDSRYL 826
 Db 1928 QTKEKCLEKONBNKOKVIVCIEELSVWTSERNQLRGELDTMSKKTTALDQLSERKEXT 1987
 Qy 827 HLLTGUFEMMINGADAVHFRUTMFLKIVKWHEDVOLFKKCCSVWYTGSISLNPNCSVK 886
 Db 1988 Q-----ELESHQSCLHCTQVAEAEVKETE-----LIOTLSSDVSELJKDKTH 2031
 Qy 887 TVLQTOALYVCAMISSQKOCKHOLA-----SISSP 918
 Db 2032 IOEKLOSLERKSQASLTLKCELENQIAQNLNEKEELVKSESLSOARLSRSDYEKUNSKA 2091
 Qy 919 VVTSLI-----INIGSPYKVR--RAAOIQLQALSGVASPFYLIDHLLISKAETISDA 970
 Db 2092 LEALUVEKGFFARLRSSTQBEVHQRLRGJEKVRVIAEKBLQ-----HIAEKLEREREN 2148
 Qy 971 AVYDQDLATLFEELOREKKU-KSHQKL-----SBLTKNLSCVSCPSYIAKOL 1018
 Db 2149 ---DSIKDKVNRLEBRELONSEENOLVILDAENSKAETVLTQKQIEEMARSLKFELD 2204
 Qy 1019 M-----KVLOGUNGBM-VISOLLPMVAEQLEKIQEKPPTAV-LKOBAM---VLHL 1062
 Db 2205 VTLRSBKENITRQIOUHQGQSLSEJLQSFKSLI-EKQAOIQIKERSTAVMLO 2262
 Qy 1063 TLGKNEFSTSILNED----PESLDIFKAVHFTKELJAGMPTQITALEKIKTPFFA 1116
 Db 2263 OLKELNEAVALCQDOBIMKATEOSLIDPPTEEHOLR-----NSIEKURARLEA 2311
 Qy 1117 AISDEKQOKLRLMFULLVCKNSHCACTVSSVKGIS-----VNAEQVRLIEPD 1169
 Db 2312 ---DEBKQKOLCUTQOLKE-----SEHHADLKGRVNRLELEARTNBHALAENSK 2362
 Qy 1170 -----KAKPLGTVOORKRORMQOKKSQDPLSVDEVGGSYWRVILIEJLQHKCKURSP 1223
 Db 2363 GEVETIKAKEGMOSIRGEL-----DVTIRS-----EKENTNLQEKRISL 2410
 Qy 1224 QILVPTLNLISRCLEPLFOBOGNMNEYTKQKLLSCLNLTICKLSPDGKPKF---TLDIE 1281
 Db 2411 BILNSSPENI---LOKBEQEVKOMKSTAMEMLOQTLKELNERVAALHINDOBACKAK 2466
 Db 1282 KENVELIQCOTRILSE-----MPQTHHALLLGTWAGF-----PDVTLHN 1323
 Qy 2467 BONLSSQVECBLERQQLQGLDDEAKNNYVLOSSVNGLQTOVEDGQKLEKKDEBISRL 2526
 Qy 1324 MISIFTFMGANWMRLLDT-----YSECVINKVWVMP-ALTQDSG---DSIEV-- 1368
 Db 2527 KNQDOEQOLVSKSVOGEVBHOLKRNQBLNLUFELWQKIOVQKQNLQDPLTLEVLQ 2586
 Qy 1369 --SRNVE--EIVVKISYFVDAHPVPE-----HRLPLVQVLYTGAEK--FL 1412
 Db 2587 SYKOLENEDBLTRDKMFPVERKOMKTAKETELOREMHEMAQKTAELQBELGEKGRLA 2646
 Qy 1413 WILLLFEEVYKTVKLASAYGEKOIALEADTEFWFSVCC-----1452

Run on: June 29, 2006, 00:41:09 ; Search time 144.903 Seconds
 Copyright (c) 1993 - 2006 Biocceleration Ltd.

OM protein - protein search, using sw model

Perfect score:	US-09-603-665-5
Score:	10833
Sequence:	MTLAAQQLQLALPQDSL.....COKTIQQLETIVGEPLOSSYF 2144 (without alignments)
Scoring table:	BLOSUM62
Gappen:	10.0 Gapext 0.5
Searched:	2589679 seqs, 457216429 residues
Total number of hits satisfying chosen parameters:	2589679
Minimum DB seq length:	0
Maximum DB seq length:	200000000

Post-processing: Minimum Match 0%
 Maximum Match 100%
 Listing first 45 summaries

Database : A_Geneseq_B;*

- 1: geneseqp1980s;*
- 2: geneseqp1990s;*
- 3: geneseqp2000s;*
- 4: geneseqp2001s;*
- 5: geneseqp2002s;*
- 6: geneseqp2003as;*
- 7: geneseqp2003bs;*
- 8: geneseqp2004s;*
- 9: geneseqp2005s;*
- 10: geneseqp2006s;*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query	Match Length	DB ID	Description
1	10799	99.9	2144	AAB85029	Aab85029 Protein e
2	10799	99.9	2144	ADQ98900	Adq98900 Antagonis
3	10791	99.9	2144	AEE02859	Aee02859 Human S-1
4	10782	99.8	2144	ADS44902	Ads44902 Human aut
5	10222	94.5	2044	ADS44907	Ads44907 Human aut
6	9853	91.2	2036	ADS44905	Ads44905 Human aut
7	9853	91.2	2036	ADS44903	Ads44903 Human aut
8	9853	91.2	2036	ADS44904	Ads44904 Human aut
9	7756.5	71.8	1569	ABG5232	Abg5232 Novel hum
10	6124.5	56.7	1229	ADQ66400	Adq66400 Novel hum
11	5026.5	46.5	1149	ADE08012	Ade08012 Novel pro
12	5025.5	46.5	1149	ADU40382	Adu40382 Novel hum
13	4506	41.7	897	ADQ96220	Adq96220 T cell ac
14	4498	41.6	897	ADQ96218	Adq96218 T cell ac
15	4494	41.6	897	ADQ96152	Adq96152 T cell ac
16	3694	34.2	734	ADM04615	Adm04615 Human pro
17	3694	34.2	734	ABC87545	Aec87545 Human cDN
18	3102	28.7	624	AEE02884	Aee02884 Protein r
19	2585	23.9	515	AAW54099	AAw54099 Homo sapi
20	2061	19.1	408	ADQ84906	Adq84906 Human aut
21	2051.5	19.0	2095	ABP85242	Abp85242 Drosophil
22	1908	19.0	2096	ADQ89606	Adq89606 Antagonis
23	1779	16.5	349	ADP24166	Adp24166 PRO polyp

ALIGNMENTS

RESULT 1		
ID	Label	Description
AAB85029	standard; protein;	protein; 2144 AA.
XX		
AC		AAB85029;
XX		
DT		06-AUG-2001 (first entry)
XX		
DB		Protein encoded by BAP28 cDNA consisting of exons 1 to 45.
XX		
KW		BAP28; prostate; tumour; cancer; diagnostic; genetic analysis.
XX		
OS		Homo sapiens.
XX		
PH		Key
FT		Misc-difference 1694
FT		/label= Ser or Asn
FT		Misc-difference 1854
FT		/label= Ala or Val
FT		Misc-difference 1967
FT		/label= Asp or Asn
FT		Misc-difference 2117
FT		/label= Gly or Glu
XX		
PN		WO200100669-A2.
XX		
PD		04-JAN-2001.
XX		
PF		23-JUN-2000; 2000WO-IB001183.
XX		
PR		25-JUN-1999; 99US-0141323P.
PR		18-JAN-2000; 2000US-0176880P.
XX		
PA		(GST) GENSET.
XX		
PT		Barry C., Bougueret L., Chumakov I., Cohen-Akenine A;
XX		
DR		WPI; 2001-367032/38.
DR		N-PSDB(AAF83909, AAF83910).
XX		
PT		New BAP28 polynucleotides and polypeptides overexpressed in prostate
PT		cancer cells for diagnosing prostate tumors, e.g. by hybridization or
PT		polymerase chain reaction assays.
XX		
PS		Claim 14; Page 297-304; 349pp; English.

The invention is directed to BAP28 polypeptides, BAP28 polynucleotide sequences and regulatory region located at the 3', and 5' ends of the BAP28 coding region. The BAP28 polypeptides can be expressed by standard recombinant methodology. BAP28 polynucleotides and polypeptides have been found to be over expressed in prostate tumour cells, therefore levels of BAP28 expression and/or activity may be assayed (e.g. by polymerase chain reaction (PCR)) to diagnose patient suffering from or susceptible to prostate cancer. Antibodies specific for the BAP28 polypeptides are useful as diagnostic reagents. Biallelic markers of the BAP28 gene are useful in genetic analysis. The present sequence represents a protein encoded by a first cDNA sequence of the BAP28 gene consisting of the exons 1 to 45

Copyright (c) 1993 - 2006 Biocceleration Ltd.

GenCore version 5.1.9

OM protein - protein search, using sw model

Run on: June 29, 2006, 00:42:00 ; Search time 51.7106 Seconds

(without alignments)
3989.395 Million cell updates/sec

Title: US-09-603-665-5
Perfect score: 1003
Sequence: 1 MTLAQLQRLALPQSDASL.....COKTIOOLETVLGERIPLQSYF 2144

Scoring table: BLOSUM62
Gppop 10.0 , Gapext 0.5

Searched: 283416 seqs, 96216763 residues

Total number of hits satisfying chosen parameters: 283416

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%

Listing first 45 summaries

Database : PIR_801,*

1: pir1:*

2: pir2:*

3: pir3:*

4: pir4:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No. Score * Query Length DB ID Description

1 990.5 9.2 1649 2 T3993B hypothetical protein SPBC23E6_04C - fission yeast (Schizosaccharomyces pombe)

2 956 8.8 1769 2 S5337B probable membrane protein

3 857 7.9 1650 2 T27864 hypothetical proteo

4 317.5 2.9 2670 2 T37919 GCN homolog - fib

5 279 2.6 2672 2 A48126 translation activa

6 272.5 2.5 2954 2 T14156 kinesin-related pr

7 268.5 2.5 2663 2 S28261 centromere protein

8 268.5 2.5 2819 2 A90551 conserved hypothe

9 261.5 2.4 1727 2 T50073 myosin-like coiled

10 250 2.4 4392 2 T29042 hypothetical prote

11 257 2.4 2297 2 AB24942 probable nuclear p

12 256.5 2.4 1837 2 T41023 utrophin - human

13 255.5 2.4 2433 1 S28381 hypothetical prote

14 254.5 2.4 2895 2 H85362 TRAP-like protein

15 252 2.3 3655 2 T38084 364K Golgi complex

30	232.5	2.2	2470	2	S57085
31	232.5	2.2	2712	2	T05113
32	231	2.1	2253	2	T3036
33	230	2.1	2335	2	T40186
34	229.5	2.1	2104	2	T38774
35	228	2.1	4131	2	T21085
36	227	2.1	1922	2	T0637
37	227	2.1	1939	2	T18372
38	226	2.1	1447	2	F82909
39	225.5	2.1	1679	2	S48385
40	225.5	2.1	3225	2	T52300
41	225.5	2.1	3259	1	A56539
42	224.5	2.1	1819	2	A71928
43	224	2.1	1959	1	A33977
44	224	2.1	2748	2	S57976
45	2.1	3738	2	T05501	

ALIGNMENTS

RESULT 1

T3993B hypothetical protein SPBC23E6_04C - fission yeast (Schizosaccharomyces pombe)

C;Species: Schizosaccharomyces pombe

C;Date: 03-Dec-1999 #sequence_revision 03-Dec-1999 #text_change 09-Jul-2004

C;Accession: T39938 R;Wood, V.; Rajandream, M.A.; Barrell, B.G.; Pohl, T.

submitted to the EMBL Data Library, May 1998

A;Reference number: 221892 A;Accession: T39938

A;Status: preliminary; translated from GB/EMBL/DBJ

A;Molecule type: DNA A;Residues: 1-1649 <WOO>

A;Cross-references: UNIPROT:060179; UNIPARC:UPI000013AF93; EMBL:AL023287; PIDN:CAA18872.1

A;Experimental source: strain 972h-; cosmid c23E6 C;Genetics:

A;Gene: SPDB:SPBC23E6_04C A;Map position: 2

Query Match 9.2%; Score 990.5; DB 2; Length 1649; Best Local Similarity 21.1%; Pid 2e-40; DB 2; Matches 469; Conservative 379; Mismatches 718; Indels 655; Gaps 78;

Qy 2 TSLAQQLQRLALPQSDASLRSDEV--ISLEDPKEAATIDRTAFAIGCTGRELGI 58

Db 3 SSLOQQLQKNTI--QSN-NVILKINKIRRAPSLLYDPKVAADMDEBIVYTAWSGFHELAVH 58

Qy 59 DPSFQPEARFLPSOALKTVERSVQTKAVNKQLDENISLFLIHLSPYFLLKPAQKCLEWL 118

Db 59 EPRLVYFEKLGLGROSQVDRVLLNRTEKIDLCVQLRLLAPPFTERNALKYLEWL 118

Qy 119 HRFHFLYMQNDLSLACAVLPYHETRIFRVTRVFLQKINNSKHWFWNLPLVKQSGVPLAKGT 178

Db 119 RRFSTHEYVSDEFILSFLPFDHFFPARLIGCSK-PKSRLPLFENAIKMP-VLISRADI 176

Qy 179 ITHCYKDGLGMDTICSLVYTKSVKVFAEYGGSSAQRLVLAFYASTIVSVAADSDN 237

Db 177 WHALSRDKEFFAMFAQVONTAESHNMYP-----ELARFWAGTMVEVIAWHSSNEDP 229

Qy 238 --ITAKLPPYIQKGK--SSLPLDYRAATMICOLQISVKMTENFVNLSASQIKLTKI 293

Db 230 NVLUDRFFLRVSVASYVSS--DFOIAGMLLSSAALPLSPSIIPPLVSAITDRLSF 287

Qy 294 PSLIKDGLSCLIVIL-----ORQKPESLG-----

Db 288 -DNKPAICVGHUQFCSEFEHQEQLKLESGASSLILLESQEHRDDEFFSYWSL 346

Qy 318 ---KKPFPHLCNPVDLITLHGSETDVSPILYRMLPHLVVSIIRHVTGEETGMG 372

Db 347 IKSRKQDKKRLISLD--TSISQIRVTHQAKFLLSVP-----WNQDFKAL- 392

Qy 373 QIYKRHLEAITLTKI-SLKNNLHLLASLLFEEYIYSQQEMDSNVSKVSLNQFLPLIRLL 432

1-phosphatidylinos
hypothetical prote
nuclear/mitotic ap
probable phosphati
myosin 3 heavy cha
hypothetical prote
hypothetical prote
repeat organelar
hypothetical prote
hypothetical prote
giantin - human
giantin - human
cag island protein
myosin heavy chain
nuclear migration
hypothetical prote

Db
QY 493 PVRILAMOHHLKKIMKTSKEGUDSFKEAVLARLGDDNIDVULSAISAFEFKEHESSEV 552
Db 471 -----LKKISEIDSOTSFEOLAYADSA-----ADLDSSVFISL----- 470
QY 553 TISNLNLFQRAELSKNGEWEYEVLKIAADILIKEETISENDOLSNQVVCUPLPFFV--IN 610
Db 493 IILSIEL-----RKTIEENKDWDYII-----LPNVLYSLQ 524
QY 611 NDTESEAENKIAJYLUSKGICSLIPLLRGWEALEVIKSTKPGKLIGVANQKMIELLAD 670
Db 525 SKOTEVRSRALNLIT-----FLELRNE 547
QY 671 NINLGDPSSMLKMDVLISUGEESEPNLKOKUTFHVLISVLVSCSSLKETHFPFAIRVF 730
Db 548 NL-----EFSIYGMDDND----- 562
QY 731 SLLQQKTKIKGLESVTAVIPSEWHTELMDRGIPBFLWAHYVELNSTORVAVEDSVFLV 790
Db 563 ---KNLRWLSPVET-----KYCCSDILLBRSSSIGI-----DCTYLF 596
QY 791 FSUKKFIVALKAPPSPPKGDIWMNPOLKEDSRDYHLLIGLFEMMLNGADAVHFRVLMK 850
Db 597 SYPERLITEKKPKNAASK-----EIAVT 619
QY 851 LFTFKVHLEDVFQOLKFCSVWMTYGSSSLNPINCNSVKTQLQTAOYVGAMMSSQTOCKH 910
Db 620 SFLSH-----AACKSLSN----- 633
QY 911 QLASSISSPPVUTSLIJNLGSPVKEVERRAAIQLOALSGVASPFYLIIDHLISKAETTSDA 970
Db 634 -----WRLLELTTRYHG-----KVED----- 651
QY 971 AVYIQDLATLTFEEYOREKUKKSHQKUSETUKNLSCVSYCSYIAKDMKVQGVNGEMV 1030
Db 652 ---AKMQTILLPRBLSEFNS-EKFKTVSKREVEALVNCFNHTS-FTSLISLFLSSNIV 704
QY 1031 LSQNLPMABOLLEKIQKOEPTEAVLKOBAMVHLILTGKYNENFSVSLUNEEDPKSUDIIFPAHV 1090
Db 705 LSOAI--CRRIVE-TQS-----HL-----HDQPRLE-FVKAVI 733
QY 1091 TXKELYAGMPTQITALEKIKTPFFRAISDEKVQOKLRLMFLDULVNCKNSHCAQTUVSS 1150
Db 734 SQDE---OPHYWYDVLDSKIP-----DTV 755
QY 1151 FKGJSVNAQBQVRFLEPPDKAKPLGTVOQKRQKOMQOKKSODLESYOBGGSYWORVTLI 1210
Db 756 FK-----KLIGSVRLVKEKPNAIKRKPIIDS-HIFQGDVQRTRI 794
QY 1211 LEHQHKKKLRSPOILVPUTPNLRCLEPLPQOEGCNMEYTKOLINSCLLN-----CQK 1265
Db 795 LELETKNASASYPKLASP-LFEVLMNSVI-ALKEDIVSSNNYLQQLLGLLYEMIGASPIE 852
QY 1266 LSPCGKPKDILDEBEKFNUBLIVQCIRLSEMPOTHHALLJLGTAGIFPKVLHIMS 1325
Db 853 LSP-----SIRIDLWLGSCIRSTNNPOIQNKALLIVSALANAAPERVLHGMP 899
QY 1326 IFTFMGANYMRLLDITYSFQVINKTVKVNIPALI-----SDSGSIEVSRVNEIWVKLISVF 1383
Db 900 IFTFMGSTVLSRDASFHSIHVEQTVKTVVISALRUGKDFDSL-----LVSCF 947
QY 1384 VDAPLPHVPERRPLPITVOLVQDUTGAKKFLWILLLIFEQYVTRVLAAYGERDALEAD 1443
QY 948 VNAPRPHIOPRRRLRYLVLWTIGSNRFLSVLI----QFAEKMLA---KSTNUVAI 998
QY 1444 TEFFFSVCCERSQHOIQSLANILQQLKLBEK-BETIPKAVSFUKSESQEMQVFN 1501

RESULT

553378 probable membrane protein YJL109c - yeast (*Saccharomyces cerevisiae*)

RESULT 2
S53378 membrane protein YJL09c - yeast (*Saccharomyces cerevisiae*)
probable protein-protein interaction

GenCore version 5.1.9
 Copyright (c) 1993 - 2006 Biocceleration Ltd.

OM protein - protein search, using sw model

Run on: June 29, 2006, 00:41:34 ; Search time 219.344 Seconds
 (without alignments)
 9041.668 Million cell updates/sec

Title: US-09-603-665-5
 Perfect score: 10803
 Sequence: 1 MTSIAQOLRLALPQSDASL.....CQKTIQOLETVLGEPLQSYF 2144

Scoring table: BL2S6M2
 Gapop 10.0 , Gapext 0.5

Searched: 2849598 seqs, 925015592 residues

Total number of hits satisfying chosen parameters: 2849598

Minimum DB seq length: 0

Maximum DB seq length: 20000000

Post-processing: Minimum Match 0%
 Maximum Match 100%
 Lasting first 45 summaries

Database : UniProt 7.2;*
 1: uniprot_sprot;*
 2: uniprot_trembl;*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query	Match length	DB ID	Description
1	10791	99.9	214	1 HEATR_HUMAN	Q9h583 homo sapien
2	10342.5	95.7	2063	2 Q5T3Q7_HUMAN	Q5t3q7 homo sapien
3	5807.5	53.8	2159	2 Q7SY41_BRADY	Q7sy48 brachyandio
4	5562	51.5	1106	2 Q6P191_HUMAN	Q6p197 homo sapien
5	4944	45.8	2288	2 Q4T7Z3_TETNG	Q4t2z3 tetraodon
6	4891	45.3	1100	2 Q3UNW7_MOUSE	Q3unw7 mus musculus
7	4729	43.8	958	1 HEATR_MACFM	Q9gm4 macaca fasci
8	4516	41.8	1090	2 Q3TNEL_MOUSE	Q3tnel mus musculus
9	4494	41.6	2077	2 Q8N7J7_HUMAN	Q8n717 homo sapien
10	4072	37.7	984	2 Q3TMK0_MOUSE	Q3tmk0 mus musculus
11	3935.5	36.4	1336	2 Q7T152_BRARE	Q7t152 brachydanio
12	3669.5	34.0	1278	2 Q7T153_BRARE	Q7t153 brachydanio
13	3315.5	30.7	743	2 Q3VIX6_MOUSE	Q3vix6 mus musculus
14	3099	28.7	733	2 Q8BLJ4_MOUSE	Q8bj4 mus musculus
15	3095	28.6	733	2 Q3TR90_MOUSE	Q3tr90 mus musculus
16	2156.5	20.0	2104	2 Q7PWD6_ANOGA	Q7pwd6 anophelidae
17	2051.5	19.0	2096	1 HEATR_DROME	Q9vn75 drosophila
18	2001.5	19.0	2096	2 Q6AW50_DROME	Q6aw50 drosophila
19	1926	17.8	408	2 Q8CC75_MOUSE	Q8cc75 mus musculus
20	1779	16.5	349	2 Q96E55_HUMAN	Q96es55 homo sapien
21	1682	15.6	349	2 Q8VCKL_MOUSE	Q8vckl mus musculus
22	1462.5	13.5	1690	2 Q8T9E7_DROME	Q8t9e7 drosophila
23	1415	13.1	2237	2 Q54ML4_DICDI	Q54ml4 dictyosteli
24	1252	11.6	2122	2 Q5CAF8_ORYZA	Q5caf8 oryza sativa
25	1194.5	11.1	1830	1 HEATR_ARATH	Q9c824 arabidopsis
26	1142	10.6	2021	2 Q5KBQ7_CRYNE	Q5kbq2 cryptococcus
27	1134	10.5	2021	2 Q5NBV7_CRYNE	Q5nbv7 cryptococcus
28	1127.5	10.4	1801	2 Q5B1X5_EMENTI	Q5b1x5 aspergillus
29	1127	10.4	1857	2 Q6BXQ5_DEBPA	Q6bxq6 debaryomyces
30	1104	10.2	1802	2 Q2ULC6_ASPRO	Q2ulc6 aspergillus
31	1088	10.1	1798	2 Q4WLI9_ASPPU	Q4wli9 aspergillus

ALIGNMENTS

RESULT 1
 HEATR_HUMAN STANDARD; PRT; 2144 AA.
 ID HEATR_HUMAN
 AC Q9h583; Q5T3Q7; Q9NW23;
 DT 01-JUN-2001, integrated into UniProtKB/Swiss-Prot.
 DT 19-JUL-2005, sequence version 3.
 DT 07-MAR-2006, entry version 33.
 DE HEAT repeat-containing protein 1 (Protein BAP28).
 GN Name=HEATR1; Synonyms=BAP28;
 DS Homo sapiens (Human);
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Buteleostomi;
 OC Mammalia; Eutheria; Earchontoglires; Primates; Catarrhini; Hominidae;
 OC Homo.
 RN [1] GLJ-2001
 RP Bouguerel et L., Chumakov I., Batty C., Cohen-Akenine A.;
 RT "A novel BAP28 gene and protein";
 RL Patent number WO0106669, 04-JAN-2001.
 RN [2]
 RP NUCBTOIDE SEQUENCE, AND VARIANTS SER-1694; ALA-1854; ASP-1967 AND
 RA [3] NCBI-TaxID=9606;
 RL Submitted (May-2005) to the EMBL/GenBank/DBJ databases.
 RN
 RNUCLEOTIDE SEQUENCE [LARGE SCALE RNA] OF 1777-2144.
 Published=14702019; DOI=10.1038/ng1285;
 RA Ota T., Suzuki Y., Nishikawa T., Otsuki T., Sugiyama T., Irie R.,
 RA Makinatsu R., Hayashi K., Sato H., Nagai K., Kimura K., Makita H.,
 RA Sakine M., Ohayashi M., Nishi T., Shibaura T., Tanaka T., Ishii S.,
 RA Yamamoto J., Saito K., Kawai Y., Isono Y., Nakamura Y., Nagahari K.,
 RA Murakami K., Yasuda T., Iwayanagi T., Wagatsuma M., Shiratori A.,
 RA Sudo H., Hosoi T., Kaku Y., Kodaira H., Kondo H., Sugawara M.,
 RA Takahashi M., Kanda K., Yokoi T., Furuya T., Kikkawa E., Omura Y.,
 RA Abe K., Kamihara K., Katsuma S., Sato K., Tanikawa M., Yanazaki M.,
 RA Niromiya K., Ishibashi T., Yamashita H., Murakawa K., Fujimori K.,
 RA Tamai H., Kimata M., Watanebe M., Hirako S., Chiba Y., Ishida S.,
 RA Ono Y., Takiguchi S., Watanebe S., Iosida M., Hotaru T., Kusano J.,
 RA Kaneko T., Takahashi-Fujii A., Hara T., Tanase T.-O., Nomura Y.,
 RA Toiya S., Komai F., Hara R., Takeuchi K., Arita M., Imose N.,
 RA Musashino K., Yukii H., Osima A., Sasaki N., Aotsuka S.,
 RA Yoshikawa Y., Matsumoto H., Ichihara T., Shionota N., Sano S.,
 RA Moriya S., Moniyama H., Satoh N., Takami S., Terashima Y., Suzuki O.,
 RA Nakagawa S., Senoh A., Mizoguchi H., Goo Y., Shimizu P., Wakabe H.,
 RA Hishigaki H., Watanebe T., Sugiyama A., Takekoto M., Kawakami B.,
 RA Yamazaki M., Watanebe K., Kumagai A., Itakura S., Fukuzumi Y.,
 RA Fujimori Y., Komiyama M., Tashiro H., Tanigami A., Fujiwara T.,
 RA Ono T., Yamada K., Fujii Y., Ozeki K., Hiroo M., Ohmori Y.,
 RA Kubabata A., Hikiji T., Kobayashi N., Inakaki H., Ikeda Y., Okamoto S.,
 RA Okitani R., Kawakami Y., Noguchi S., Itoh T., Shigeno K., Seoba T.,
 RA Matsunaga K., Nakajima Y., Mizuno T., Morinaga M., Sasaki M.,
 RA Toshimura T., Ogawa M., Hata H., Watanebe M., Komatsu T.,
 RA Mizushima-Sugano J., Satoh T., Shirai Y., Takahashi Y., Nakashita K.,
 RA Okumura K., Nagase T., Nomura N., Kikuchi H., Masuno Y., Yamashita R.,

RA Nakai K., Yada T., Nakamura Y., Ohara O., Isogai T., Sugano S.-i., "Complete sequencing and characterization of 21,243 full-length human cDNAs," Nat. Genet. 36:40-45 (2004).
 RL IDENTIFICATION BY MASS SPECTROMETRY, AND SUBCELLULAR LOCATION.
 RP MEDLINE=22317277; PubMed=12429849; DOI=10.1093/nar/gkz027; RQ
 RN Scherl A., Coute Y., Deon C., Calle A., Kindbeiter K., Sanchez J.-C., Greco A., Hochstrasser D.F., Diaz J.-J., "Functional proteomic analysis of human nucleolus.," Mol. Biol. Cell 13:4100-4109 (2002).
 RL -!- INTERACTION:
 CC Q8W71; IKNL1; NBEXP1; IntAct:EBI-373098, EBI-372406;
 CC -!- SUBCELLULAR LOCATION: Nuclear; nucleolar.
 CC -!- SIMILARITY: Belongs to the HEATR1 family.
 CC -!- SIMILARITY: Contains 1 HEAT repeat.
 CC Copyrighted by the UniProt Consortium, see <http://www.uniprot.org/terms>
 CC Distributed under the Creative Commons Attribution-NoDerivs License
 CC
 DR EMBL; AX067150; CAC2676_1; -, Unassigned DNA.
 DR EMBL; AL359321; CAA13775_1; -, Genomic DNA.
 DR EMBL; AK001221; BRA91564_1; ALT_INIT; mRNA.
 DR INTACT: Q9HS83;
 DR SWISS-2DPAGE; Q9HS83; HUMAN.
 DR Ensembl; ENSG00000119255; Homo sapiens.
 DR HGNC; HGNC:25517; HEATR1.
 DR GO:0005515; F-protein binding; IPI.
 DR InterPro; IPR012554; BP28_C.
 DR InterPro; IPR00357; HEAT.
 DR Pfam; PF08146; BP28CT; 1.
 DR Pfam; PF0285; HEAT; 4.
 DR PROSITE; PS0077; HEAT_REPEAT; FALSE_NEG.
 KW Nucleic protein; Polymorphism.
 FT CHAIN 1 2144
 FT REPEAT 2106 2142
 FT VARIANT 1694 1694
 FT VARIANT 1854 1854
 FT VARIANT 1967 1967
 FT VARIANT 2017 2017
 FT CONFLICT 348 348
 FT CONFLICT 607 607
 SQ SQUENCE 2144 AA; 242370 MW; 75908EA1754FF390 CRC64;
 Query Match 99.9%; Score 10791; DB 1; Length 2144;
 Best Local Similarity 99.7%; Pred: No: 0;
 Matches 2138; Conservative 1; Mismatches 5; Indels 0; Gaps 0;
 QY 1 MTSIAQQQLRALPQDASLILRSRDEAVASLIFPDPEKATIDRTAARRIGTCLELLGIDP 60
 Db 1 MTSIAQQQLRALPQDASLILRSRDEAVASLIFPDPEKATIDRTAARRIGTCLELLGIDP 60
 QY 61 SFEGQFEAPLFSQALKTLERSVOTKAVNKGOLDENISLPLHLSPFLKPKAQCKEFLWHR 120
 Db 61 SFEGQFEAPLFSQALKTLERSVOTKAVNKGOLDENISLPLHLSPFLKPKAQCKEFLWHR 120
 QY 121 FHILYNODSLIACVLPYHETRIVFVOLKINKNSKRHWFLVKQSVPLAKGLIT 180
 Db 121 FHILYNODSLIACVLPYHETRIVFVOLKINKNSKRHWFLVKQSVPLAKGLIT 180
 QY 181 HCYKDGLGENDFICSLVTKVKEFVPGSSAQRLVILAFASTVSLVAEDSDNIIA 240
 Db 181 HCYKDGLGENDFICSLVTKVKEFVPGSSAQRLVILAFASTVSLVAEDSDNIIA 240
 QY 1321 KLFYYIQGIKSSIPDYRAATYMICQISKVMTENYFNSLASQIKUTKISLIKDG 300
 Db 1321 KLFYYIQGIKSSIPDYRAATYMICQISKVMTENYFNSLASQIKUTKISLIKDG 300
 QY 241 KLFYYIQGIKSSIPDYRAATYMICQISKVMTENYFNSLASQIKUTKISLIKDG 300
 Db 241 KLFYYIQGIKSSIPDYRAATYMICQISKVMTENYFNSLASQIKUTKISLIKDG 300

Copyright (c) 1993 - 2005 Biocceleration Ltd.

GenCore version 5.1.9

Run on: June 29, 2006, 00:46:09 ; Search time 37.5044 Seconds

(without alignments)

5003.834 Million cell updates/sec

OM protein - protein search, using sw model

Title:

US-09-603-665-5

Perfect score:

10003

Sequence: 1 MTSLAQOLRILALPOSDSL.....CQKTIQOLETIVLGERPLQSYP 2144

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 650591 seqs, 87530628 residues

Total number of hits satisfying chosen parameters: 650591

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Lasting first 45 summaries

,

,

,

,

,

,

,

,

,

,

,

,

,

,

,

,

,

,

,

,

,

,

,

,

,

,

,

,

,

,

,

,

,

,

,

,

,

,

,

,

,

,

,

,

,

,

,

,

,

,

,

,

,

,

,

,

,

,

,

,

,

,

,

,

,

,

,

,

,

,

,

,

,

,

,

,

,

,

,

,

,

,

,

,

,

,

,

,

,

,

,

,

,

,

,

,

,

,

,

,

,

,

,

,

,

,

,

,

,

,

,

,

,

,

,

,

,

,

,

,

,

,

,

,

,

,

,

,

,

,

,

,

,

,

,

,

,

,

,

,

,

,

,

,

,

,

,

,

,

,

,

,

,

,

,

,

,

,

,

,

,

,

,

,

,

,

,

,

,

,

,

,

,

,

,

,

,

,

,

,

,

,

,

,

,

,

,

,

,

,

,

,

,

,

,

,

,

,

,

,

,

,

,

,

,

,

,

,

,

,

,

,

,

,

,

,

,

,

,

,

,

,

,

,

,

,

,

,

,

,

,

,

,

,

,

,

,

,

,

,

,

,

,

,

,

,

,

,

,

,

,

,

,

,

,

,

,

,

,

,

,

,

,

,

,

,

,

,

,

,

,

,

,

,

,

,

,

,

,

,

,

,

,

,

,

,

,

,

,

,

,

,

,

,

,

,

,

QY 1758 L---SALAALQKVENTLPHFISPVYEGILSOVIRHKBITSMGSA----- 1799
 Db 341 LLQGSVIVLVLSCYIKKIPAMSTTLBAVLTLSLDIDNHRSSVLDLIVDHMDLAOV 400
 Qy 1800 -SQANIRLTSLLKTTATTLPAPRVLPAIKKYKQEAKWNKHMGPMPSILOEHGQXMKKE 1058
 Db 401 KSLCHWMLT-----KKFY--TNDNSGNIGFLKTQATINRMEEK 438
 Qy 1859 ELTSHQSQLTAFFLEALDFRAQHSEN-----DLEVGKTENCIDCLVAMVUKSEVFR 1913
 :|: ;|: ;|: ;|: ;|: ;|: ;|: ;|: ;|: ;|: ;|: ;|: ;|: ;|: ;|: ;|: ;|:
 439 QATTOTLFLMRWLISAFER-QYSEDNDNKFDDNTTHRLESSFHGCAIAFWMKNDKSFR 497
 :|: ;|: ;|: ;|: ;|: ;|: ;|: ;|: ;|: ;|: ;|: ;|: ;|: ;|: ;|: ;|:
 Db 1914 PLFFKLFDWA-----KTEDAPDRLLTFYLNADCIACBKLGIFTLEAGHLVMPFA DT 1965
 :|: ;|: ;|: ;|: ;|: ;|: ;|: ;|: ;|: ;|: ;|: ;|: ;|: ;|: ;|:
 498 PLFANLVRWAVGEATLKNEV-SRLIAFPFRPNKQLOBLKSITSFSYLAADPTSL 555
 Qy 1966 LXQVNISKUTDQAFFESENDEPKCCLLQFTNLNCUCLKIFLDTOHTISK-RAXAMMPJY 2024
 L---KRFSEGGLSLWATN-----LRRITLGLGSSFSKYDODDWSQQRFDSCPSL 603
 Db 2025 DOLENRGGEBEKFQRVTKHLIPICTAQFSVAMADSLWKLNEYQI--LIKTRSSPKR 2081
 :|: ;|: ;|: ;|: ;|: ;|: ;|: ;|: ;|: ;|: ;|: ;|: ;|: ;|:
 604 SOLSN-----IEISIGKVLYKSVSTFVTDVSSENEYTLVHELIKYISNANENSAATK 656
 Qy 2082 FAALITVLALAKLUENYVILPESIPIFLBLMEDECEVEHOCOK-TQOLETVLGEPL 2140
 :|: ;|: ;|: ;|: ;|: ;|: ;|: ;|: ;|: ;|: ;|: ;|: ;|:
 Db 657 IWSIRTLKTFIQKOMERQWLSYLPTVYABLEDDEVEEMVERGLVRVLENVLGEPL 715
 Qy 2141 QSY 2143
 Db 717 DRY 719

RESULT 2
 US-09-248-796A-15108
 Sequence 15108 Application US/09248796A
 Patent No. 6747137

GENERAL INFORMATION:

APPLICANT: Keith Weinstock et al

TITLE OF INVENTION: NUCLEAR ACID AND AMINO ACID SEQUENCES RELATING TO CANDIDA ALBICAN FILE REFERENCE: 107196.132

CURRENT APPLICATION NUMBER: US/09/248, 796A

PRIOR APPLICATION NUMBER: US 60/074, 725

PRIOR FILING DATE: 1998-02-13

NUMBER OF SEQ ID NOS: 28208

SEQ ID NO 15108

LENGTH: 2954

TYPE: PRT

ORGANISM: Xenopus sp.

FEATURE: OTHER INFORMATION: Xenopus centromere-associated protein-E (XCENP-E) OTHER INFORMATION: member of the kinesin superfamily of microtubule other proteins

FEATURE: SOFTWARE: Patentin Ver. 2.0

SBQ ID NO 1

NAME/KEY: DOMAIN

LOCATION: (473)-(2752)

OTHER INFORMATION: rod domain

FEATURE: NAME/KEY: DOMAIN

LOCATION: (2753)-(2954)

NAME/KEY: DOMAIN

LOCATION: (1)-(472)

OTHER INFORMATION: kinesin like motor domain

FEATURE: NAME/KEY: DOMAIN

LOCATION: (473)-(2752)

OTHER INFORMATION: rod domain

FEATURE: NAME/KEY: DOMAIN

LOCATION: (2753)-(2954)

NAME/KEY: DOMAIN

LOCATION: (1)-(472)

OTHER INFORMATION: tail domain

US-09-150-867-1

Query Match 2.5%; Score 272.5; DB 2; Length 2954;
 Best local Similarity 18.7%; Pred. No. 3.1e-12;
 Matches 415; Conservative 365; Mismatches 807; Indels 529; Gaps 96;

Qy 1 MTSLAQQQLRMLAPQSDASLRSDEVAILFDPKEAATIDRPAFATGCTGABBLGIDP 60
 Db 702 MAEKONALBELA-----LMRDNFNDIL--ENETLKREI-----DLERSLKENQ 744
 Qy 61 SFEGEPEALPSQALKTLERSVQKAVN-----KOLDENISLELJHLSPYF----LLRPA 110
 Db 745 ETNEFEI----LEKETOKHEAQHQLIETHGSLKLVENAEVNQNLSEDELTKTKLKEQ 799
 Qy 111 QKCLEWL-----IHRFHLYNODSLACVLPYHETRIFVURVQOLKINNSHRW 160
 Db 840 -----KQS----LSDAEAVTRIAQKESFLRSBNLEKEKMDTSNWNOKEKAASUFK 890
 Qy 53 EELQGIDPSPEOFEAFLPSQALKTLERSVQKAVNQKDENSILFLHISPYFLKPAOK 112
 Db 60 EDLGELDSRENFKKHTLFSTSVDNLOTKDVSQDGNIDAETULGVPPYGISSIK 119
 Qy 113 CLEMWIHREHILHYMDSLACVLYHETRIFVURVQOLKINNSHRWFLPVR-QSY 171
 Db 120 AVEWLVVRREHANIHMAELMTALPYFOHVFVVEVLVNPQNLQFQIWEMLVGVYQDKLT 179
 172 PLAKGTLITHCYKUDGFMD 191

RESULT 3
 US-09-150-867-1
 Sequence 1, Application US/09150867
 Patent No. 6645748

GENERAL INFORMATION:

APPLICANT: Wood, Kenneth W.

APPLICANT: Sakowicz, Roman

APPLICANT: Goldstein, Lawrence S.B.

APPLICANT: Cleveland, Don W.

APPLICANT: The Regents of the University of California

TITLE OF INVENTION: End-Directed Microtubule Motor Required for File Reference: 18557C-000110S

CURRENT APPLICATION NUMBER: US/09/150, 867

CURRENT FILING DATE: 1998-09-10

EARLIER APPLICATION NUMBER: US 60/058, 645

EARLIER FILING DATE: 1997-09-11

NUMBER OF SEQ ID NOS: 11

SOFTWARE: Patentin Ver. 2.0

LENGTH: 2954

TYPE: PRT

ORGANISM: Xenopus sp.

FEATURE: OTHER INFORMATION: Xenopus centromere-associated protein-E (XCENP-E) OTHER INFORMATION: member of the kinesin superfamily of microtubule other proteins

FEATURE: NAME/KEY: DOMAIN

LOCATION: (1)-(472)

OTHER INFORMATION: kinesin like motor domain

FEATURE: NAME/KEY: DOMAIN

LOCATION: (473)-(2752)

OTHER INFORMATION: rod domain

FEATURE: NAME/KEY: DOMAIN

LOCATION: (2753)-(2954)

NAME/KEY: DOMAIN

LOCATION: (1)-(472)

OTHER INFORMATION: tail domain

US-09-150-867-1

Query Match 2.5%; Score 272.5; DB 2; Length 2954;
 Best local Similarity 18.7%; Pred. No. 3.1e-12;
 Matches 415; Conservative 365; Mismatches 807; Indels 529; Gaps 96;

Qy 1 MTSLAQQQLRMLAPQSDASLRSDEVAILFDPKEAATIDRPAFATGCTGABBLGIDP 60
 Db 702 MAEKONALBELA-----LMRDNFNDIL--ENETLKREI-----DLERSLKENQ 744
 Qy 61 SFEGEPEALPSQALKTLERSVQKAVN-----KOLDENISLELJHLSPYF----LLRPA 110
 Db 745 ETNEFEI----LEKETOKHEAQHQLIETHGSLKLVENAEVNQNLSEDELTKTKLKEQ 799
 Qy 111 QKCLEWL-----IHRFHLYNODSLACVLPYHETRIFVURVQOLKINNSHRW 160
 Db 840 -----KQS----LSDAEAVTRIAQKESFLRSBNLEKEKMDTSNWNOKEKAASUFK 890
 Qy 53 EELQGIDPSPEOFEAFLPSQALKTLERSVQKAVNQKDENSILFLHISPYFLKPAOK 112
 Db 60 EDLGELDSRENFKKHTLFSTSVDNLOTKDVSQDGNIDAETULGVPPYGISSIK 119
 Qy 113 CLEMWIHREHILHYMDSLACVLYHETRIFVURVQOLKINNSHRWFLPVR-QSY 171
 Db 120 AVEWLVVRREHANIHMAELMTALPYFOHVFVVEVLVNPQNLQFQIWEMLVGVYQDKLT 179
 172 PLAKGTLITHCYKUDGFMD 191

Qy 951 KALEBNALNEVTCLESYKFLNE- VECLNQOISKASEEMILKOGEBSASII--S 1005

Db 891 OLETEKSNKKMADLQELQSAFNEINYLNGLAGVKPRDLISRVELEKSEPSKOLE 950

Qy 252 SISPDYRATYMICOLISVKTMENTFNSLASSOIKLTKPSLKGSLCIVLQRO 311

GenCore version 5.1.9
Copyright (c) 1993 - 2006 Biocceleration Ltd.

Om protein - protein search, using sw model

Run on:

June 29, 2006, 00:52:34 ; Search time 134.107 Seconds

(without alignments)

7405.549 Million cell updates/sec

Title: US-09-603-665-5

Perfect score: 10803

Sequence: 1 MTLAQLQRLALPQSDASL.....CQKTIQOLETVLGEPLQSYF 2144

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched:

2037797 seqs, 463214858 residues

Total number of hits satisfying chosen parameters: 2037797

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Published Applications AA Main:*

1: /EMC_Celerra_SIDS3ptodata/2/pubpbaa/US07_PUBCOMB.pep:*

2: /EMC_Celerra_SIDS3ptodata/2/pubpbaa/US08_PUBCOMB.pep:*

3: /EMC_Celerra_SIDS3ptodata/2/pubpbaa/US09_PUBCOMB.pep:*

4: /EMC_Celerra_SIDS3ptodata/2/pubpbaa/US10A_PUBCOMB.pep:*

5: /EMC_Celerra_SIDS3ptodata/2/pubpbaa/US10B_PUBCOMB.pep:*

6: /EMC_Celerra_SIDS3ptodata/2/pubpbaa/US11_PUBCOMB.pep:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

*

Result No. Score Query Length DB ID Description

Result No.	Score	Query	Length	DB	ID	Description
1	10799	99.9	2144	5	US-10-745-237-230	Sequence 230, App
2	10782	99.8	2144	6	US-11-124-368A-277	Sequence 230, App
3	10222	94.6	2044	6	US-11-124-368A-278	Sequence 278, App
4	9853	91.2	2036	6	US-11-124-368A-276	Sequence 276, App
5	9853	91.2	2036	6	US-11-124-368A-280	Sequence 280, App
6	9853	91.2	2036	6	US-11-124-368A-281	Sequence 281, App
7	7755.8	1569	1718	5	US-10-450-763-45591	Sequence 45591, A
8	5026.5	146.5	1718	5	US-10-450-763-45591	Sequence 45591, A
9	4494	897	4	US-10-764-425-177	Sequence 177, App	
10	3694	734	4	US-10-260-3300	Sequence 3300, App	
11	2061	19.1	4	US-11-124-368A-279	Sequence 279, App	
12	2031.5	19.0	4	US-10-745-237-36	Sequence 36, Appli	
13	2051.5	19.0	4	US-11-097-142-22518	Sequence 22518, A	
14	1603	14.8	325	5	US-10-128-558-358	Sequence 358, App
15	1169	10.8	2122	4	US-10-437-963-189782	Sequence 189782, App
16	1086	10.1	1798	4	US-10-128-714-8388	Sequence 8388, App
17	1085	10.0	1814	4	US-10-128-714-3388	Sequence 3388, App
18	1056.5	9.8	1818	4	US-10-022-585-7615	Sequence 7615, App
19	673	6.2	952	4	US-10-424-590-24811	Sequence 24811, App
20	663	6.1	135	4	US-10-224-042-2412	Sequence 2412, App
21	512.5	4.7	463	4	US-10-225-115-29784	Sequence 295784, App
22	434	4.0	437	4	US-10-425-114-44447	Sequence 44447, A
23	379	3.5	77	3	US-09-304-761-37036	Sequence 37036, A
24	309.5	2.9	2228	5	US-10-511-036-2	Sequence 2, Appli
25	209.5	2.9	2230	5	US-10-511-096-4	Sequence 4, Appli
26	308	2.9	2250	5	US-10-511-096-6	Sequence 6, Appli
27	2.9	2252	5	US-10-511-096-8	Sequence 8, Appli	

ALIGNMENTS

RESULT 1

US-10-745-237-230

; Sequence 230, Application US/10745237

; GENERAL INFORMATION:

; APPLICANT: Cyclacel Limited

; APPLICANT: Glover, David

; APPLICANT: Bell, Graham

; APPLICANT: Frenz, Lisa

; APPLICANT: Midgley, Carol

; TITLE OF INVENTION: Cell-Cycle Progression Proteins

; FILE REFERENCE: P015819MO CYK CURRENT APPLICATION NUMBER: US/10/745, 237

; CURRENT FILING DATE: 2003-12-23

; PRIOR APPLICATION NUMBER: US 60/439, 123

; PRIOR FILING DATE: 2003-01-10

; PRIOR APPLICATION NUMBER: US 60/468, 492

; PRIOR FILING DATE: 2003-05-06

; NUMBER OF SEQ ID NOS: 600

; SOFTWARE: PatentIn version 3.1

; SEQ ID NO: 230 LENGTH: 2144 TYPE: PTX

; ORGANISM: Homo sapiens

; FEATURE: PEPTIDE

; OTHER INFORMATION: Q9H583

US-10-745-237-230

Query Match

Best Local Similarity 99.9% Score 10799; DB 5; Length 2144; Matches 2140; Conservative 99.8%; Pred. No. 0; Mismatches 4; Indels 0; Gaps 0;

Matches 2140; Conservative 99.8%; Pred. No. 0; Mismatches 4; Indels 0; Gaps 0;

Sequence 2140, App

Sequence 3300, App

Sequence 278, App

Sequence 280, App

Sequence 281, App

Sequence 45591, A

Sequence 177, App

Sequence 3388, App

Sequence 279, App

Sequence 36, Appli

Sequence 22518, A

Sequence 24811, App

Sequence 189782, App

Sequence 8388, App

Sequence 3388, App

Sequence 7615, App

Sequence 24811, App

Sequence 2412, App

Sequence 295784, App

Sequence 44447, A

Sequence 37036, A

Sequence 2, Appli

Sequence 4, Appli

Sequence 6, Appli

Sequence 8, Appli

Sequence 215, App

Sequence 36864, A

Sequence 256711,

Sequence 214, App

Sequence 1, Appli

Sequence 749, App

Sequence 9, Appli

Sequence 9003, Ap

Sequence 10685,

Sequence 11, Appli

Sequence 2, Appli

Sequence 3, Appli

Sequence 20, Appli

Sequence 1, Appli

Sequence 41, Appli

Sequence 2, Appli

Sequence 338, App

Db 241 KLFPPYIQQKLKSLPDYRAATYMICQISVKVTMENTFVNLSASQIKLTKPLSIKG 300 Qy 1381 SYFDALPHVPEHRRPILVOLVDTLGAEKFTWILLLFCSYTCKVLAAYGEKDAIL 1440
 301 LSCUUVLORQKESELGKPPFLCNVDLITLHGSETYDVSPLRQYMLPHVWITH 360 Db 1381 SYFDALPHVPEHRRPILVOLVDTLGAEKFTWILLLFCSYTCKVLAAYGEKDAIL 1440
 301 LSCUUVLORQKESELGKPPFLCNVDLITLHGSETYDVSPLRQYMLPHVWITH 360 Qy 1441 EADTEFWSVCCBFSVQHOIQSLMNLQYLKLPEEKEETPKAVSFNKSQEMLQVF 1500
 361 HVTGEGETECDGQIYKRHLAULTKISLKNLDLHLASLFFEYISYSQEQEDNSKVL 420 Db 1441 EADTEFWSVCCBFSVQHOIQSLMNLQYLKLPEEKEETPKAVSFNKSQEMLQVF 1500
 421 LNSOFLPLRLLESKPYPTUDVVLREBEHLKIAIDLKQELPHOFVSLSTSGGKQPLADSD 480 Db 1501 NVEHTTKQLRKPFLSVFMSQOLLSSNNPLKKWESGGPPLKGLEBLTUTVYISA 1560
 481 TSLMSLNHLAPTRILAMMHKKMKTSKEGVDSFKEAVLRLGDNDIVVLSA 540 Qy 1561 VAOQMERNAKDLTVKWRALLSKAYDLDKVALLPBTFIVIRGLVGNPLPSVRKAL 1620
 481 TSLMSLNHLAPTRILAMMHKKMKTSKEGVDSFKEAVLRLGDNDIVVLSA 540 Db 1561 VAOQMERNAKDLTVKWRALLSKAYDLDKVALLPBTFIVIRGLVGNPLPSVRKAL 1620
 541 FEIFKEHFSSEVTISNLNFQARELSNGEWEYKIAIDLKEIENDOLSNOV 600 Db 1621 DLANKQONISUKKTVTREKUNPLATVORKKEGBEOAINROTALYTKLCKN 1680
 541 FEIFKEHFSSEVTISNLNFQARELSNGEWEYKIAIDLKEIENDOLSNOV 600 Qy 1621 DLANKQONISUKKTVTREKUNPLATVORKKEGBEOAINROTALYTKLCKN 1680
 601 VCLJLPVVINDDTSAEMKIAIVSKSGCISLHPLLRGWEAENLVISTKPGKIGA 660 Db 1681 FGAENDPDPVPLVXTAVKLIAPERKEEONVSLGALCTAEVSTLEAIPOLPSLML 1740
 601 VCLJLPVVINDDTSAEMKIAIVSKSGCISLHPLLRGWEAENLVISTKPGKIGA 660 Qy 1681 FGAENDPDPVPLVXTAVKLIAPERKEEONVSLGALCTAEVSTLEAIPOLPSLML 1740
 661 NOKMELADDNINQDPSMUKMVNDLISYGEEFNKLQYTFVILSUVSCSSLKE 720 Db 1741 LTWMNTSLSVSEYVLLSALAQLQVETLPHRISPYLEGISQVHETKSEMGAS 1800
 661 NOKMELADDNINQDPSMUKMVNDLISYGEEFNKLQYTFVILSUVSCSSLKE 720 Qy 1741 LTWMNTSLSVSEYVLLSALAQLQVETLPHRISPYLEGISQVHETKSEMGAS 1800
 721 THFPPAIRVSLUOKKIKLUESVITAVEIPSEWHTBLMDRGIPVBLWAHYVEENSTOR 780 Db 1801 QANRILTSLSKTTTLPVLLPAIKTYQIKRNWKHMMGPMSLOEHIGMKBEL 1860
 721 THFPPAIRVSLUOKKIKLUESVITAVEIPSEWHTBLMDRGIPVBLWAHYVEENSTOR 780 Qy 1801 QANRILTSLSKTTTLPVLLPAIKTYQIKRNWKHMMGPMSLOEHIGMKBEL 1860
 781 VAWESVFLVSLKKEIYALGAKPSPKGDWNPOLKEDSRDYLHLIGLPEMLNGA 840 Db 1861 TSHQSLTAAFLTAFLDFAQHSENDLEBEVGKTCIDCVAMVTLSEVTFRFLPKLF 1920
 781 VAWESVFLVSLKKEIYALGAKPSPKGDWNPOLKEDSRDYLHLIGLPEMLNGA 840 Qy 1861 TSHQSLTAAFLTAFLDFAQHSENDLEBEVGKTCIDCVAMVTLSEVTFRFLPKLF 1920
 Db 841 DAVHFRVLMKUFVHEDVOLFKECSVLMYGSISLNPJCSVKTVOQTLQYGCAM 900 Db 1921 DWAKTEDAPKDRLLTFLYNLADCIAEKGFLFTAGHLYKPFAPTDXQVNSISKDEAFFD 1980
 Qy 901 LSSQKTOCKHOLASISSPVVTSLLINLGPSPKEVRAIOCLQALSGVASPFYLIDHLI 960 Db 1921 DWAKTEDAPKDRLLTFLYNLADCIAEKGFLFTAGHLYKPFAPTDXQVNSISKDEAFFD 1980
 901 LSSQKTOCKHOLASISSPVVTSLLINLGPSPKEVRAIOCLQALSGVASPFYLIDHLI 960 Qy 1981 SENDPERCCULLQFLINCLYKIFLTDQTHQISKERAAXALMMPVQOLENBLIGGEEKFOER 2040
 841 DAVHFRVLMKUFVHEDVOLFKECSVLMYGSISLNPJCSVKTVOQTLQYGCAM 900 Db 2041 VTKHLIPICTAOFSVANADDSWKPLNYQILKTRDSSPKURFALITVLAKEKUNYI 2100
 961 SKABEITSDAAYVIQDIALTFLERKEKKLHSQKUSSETKOLICSCYSCSYIADLMK 1020 Db 2041 VTKHLIPICTAOFSVANADDSWKPLNYQILKTRDSSPKURFALITVLAKEKUNYI 2100
 961 SKABEITSDAAYVIQDIALTFLERKEKKLHSQKUSSETKOLICSCYSCSYIADLMK 1020 Qy 2101 VLPPESTPFLARLMEDECEEVHQCTKIQOLETVIGEPLSYF 2144
 Db 1021 VIQGVNGEMWUSQOLPMAQELEKIQKEPTAVKDEAMVLHUTLGKNEFSLNEDPK 1080 Db 2101 VLPPESTPFLARLMEDECEEVHQCTKIQOLETVIGEPLSYF 2144
 Db 1021 VIQGVNGEMWUSQOLPMAQELEKIQKEPTAVKDEAMVLHUTLGKNEFSLNEDPK 1080 Qy 2101 VLPPESTPFLARLMEDECEEVHQCTKIQOLETVIGEPLSYF 2144
 1081 SLDFIKAVHTTKELYAGMPTQITALEKITKOPFAATSDEKVQKULRMFLDVNCN 1140 RESULT 2
 1081 SLDFIKAVHTTKELYAGMPTQITALEKITKOPFAATSDEKVQKULRMFLDVNCN 1140 US-11-124-368A-277
 1081 SLDFIKAVHTTKELYAGMPTQITALEKITKOPFAATSDEKVQKULRMFLDVNCN 1140 Sequence 277, Application US/11124368A
 Db Publication No. US20050287559A1
 1141 SHQQTVSFVKGJSVNAQVRLELEPPDKAOKGTVOQKROQKQKSDQESVQBG 1200 GENERAL INFORMATION:
 1141 SHQQTVSFVKGJSVNAQVRLELEPPDKAOKGTVOQKROQKQKSDQESVQBG 1200 APPLICANT: Michele Cargill
 Qy APPLICANT: James J. Devlin
 Db APPLICANT: May Luke
 Qy TITLE OF INVENTION: Genetic Polymorphisms Associated with
 Title of Invention: Vascular Diseases, Methods of Detection and Uses Thereof
 Db TITLE OF INVENTION: Vascular Diseases, Methods of Detection and Uses Thereof
 1201 GSYWQRVTILELJLQHKKQLSPQILVPLTNLSCLEPLPQOGNMNMYTKOLILSC 1260 FILE REFERENCE: C101524
 1201 GSYWQRVTILELJLQHKKQLSPQILVPLTNLSCLEPLPQOGNMNMYTKOLILSC 1260 CURRENT APPLICATION NUMBER: US/11/124,368A
 1261 NICOKLSPDGKIKRDIDBEKFENELVOCIRISBEMPOTHHALLJLGTAGIFPDKV 1320 CURRENT FILING DATE: 2005-05-09
 1261 NICOKLSPDGKIKRDIDBEKFENELVOCIRISBEMPOTHHALLJLGTAGIFPDKV 1320 PRIORITY NUMBER: US 60/1568, 845
 1261 NICOKLSPDGKIKRDIDBEKFENELVOCIRISBEMPOTHHALLJLGTAGIFPDKV 1320 PRIORITY FILING DATE: 2004-05-07
 1321 HNIMISIFTMGANTNRDITYSFQVINKTUVNIPALIOSDGSIEVERVERBIVKII 1380 PRIORITY APPLICATION NUMBER: US 60/625, 936
 1321 HNIMISIFTMGANTNRDITYSFQVINKTUVNIPALIOSDGSIEVERVERBIVKII 1380 NUMBER OF SEQ ID NOS: 2112
 Db SOFTWARE: FastSEQ for Windows Version 4.0

Copyright (c) 1993 - 2006 Biocceleration Ltd.

Gencore version 5.1.9

Om protein - protein search, using sw model

Run on: June 29, 2006, 00:53:49 ; Search time 11.932 Seconds

Sequence: (without alignments)
4243.670 Million cell updates/sec

Title: US-09-603-665-5

Perfect score: 10803

1 MTSLAQQLQRUALPOSDSL.....CQKTIQOLETVLGEPLQSYF 2144

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 103426 seqs, 23619683 residues

Total number of hits satisfying chosen parameters: 103426

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Published Applications_AA_New:*

1: /EMC_Celerra_SIDS3_ptodata/2/pubpaa/US09_NEW_PUB.pep:*

2: /EMC_Celerra_SIDS3_ptodata/2/pubpaa/US06_NEW_PUB.pep:*

3: /EMC_Celerra_SIDS3_ptodata/2/pubpaa/US07_NEW_PUB.pep:*

4: /EMC_Celerra_SIDS3_ptodata/2/pubpaa/US08_NEW_PUB.pep:*

5: /EMC_Celerra_SIDS3_ptodata/2/pubpaa/US09_NEW_PUB.pep:*

6: /EMC_Celerra_SIDS3_ptodata/2/pubpaa/US10_NEW_PUB.pep:*

7: /EMC_Celerra_SIDS3_ptodata/2/pubpaa/US01_NEW_PUB.pep:*

8: /EMC_Celerra_SIDS3_ptodata/2/pubpaa/US06_NEW_PUB.pep:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match length	DB ID	Description
1	3694	34.2	7	US-11-293-697-3300
2	235	2.2	3113	6 US-10-505-928-325
3	207.5	1.9	2209	7 US-11-301-554-1903
4	205	1.9	93	6 US-10-449-902-34905
5	188.5	1.7	1328	6 US-10-504-973-32
6	184.5	1.7	9535	6 US-10-471-571A-4496
7	179	1.7	1206	7 US-10-449-902-43305
8	175	1.6	1205	7 US-11-293-697-2967
9	175	1.6	1842	6 US-10-511-937-2929
10	173.5	1.6	1176	6 US-10-449-902-46322
11	173.5	1.6	1325	6 US-10-449-902-51794
12	166	1.5	1346	6 US-10-449-902-55021
13	163.5	1.5	2871	6 US-10-505-928-100
14	161	1.5	1281	6 US-11-449-902-47146
15	159.5	1.5	2671	6 US-10-505-928-784
16	158.5	1.5	1116	6 US-10-449-902-7106
17	157.5	1.5	980	7 US-11-202-296
18	155	1.4	764	6 US-10-449-902-45961
19	152.5	1.4	2870	6 US-11-289-102-378
20	152	1.4	748	6 US-11-449-902-53645
21	152	1.4	2354	7 US-11-289-102-242
22	151.5	1.4	1368	7 US-11-283-329-666
23	151.5	1.4	1120	6 US-10-505-928-213
24	150	1.4	1050	6 US-10-471-571A-4624

ALIGMENTS

RESULT 1
US-11-293-697-3300
; Sequence 3300, Application US/11293697
; Publication No. US20060105376A1

GENERAL INFORMATION:

; APPLICANT: HELIX RESEARCH INSTITUTE
; TITLE OF INVENTION: Novel full length cDNA
; FILE REFERENCE: H1-A016
; CURRENT APPLICATION NUMBER: US/11/293, 697
; CURRENT FILING DATE: 2005-12-05
; PRIORITY APPLICATION NUMBER: US/10/108, 260
; PRIORITY FILING DATE: 2002-03-28
; NUMBER OF SEQ ID NOS: 5458
; SOFTWARE: PatentIn Ver. 2.1
; SBQ ID NO: 3300
; LENGTH: 734
; TYPE: PRT
; ORGANISM: Homo sapiens

US-11-293-697-3300

Query Match 34.2%; Score 3694; DB 7; Length 734;
Best Local Similarity 100.0%; Pred. No. 7; se-206;
Matches 734; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY Sequence 3300, AP

QY Sequence 325, AP

QY Sequence 1903, AP

Db Sequence 34905, A

QY Sequence 35765, A

QY Sequence 32, APP

QY Sequence 4496, AP

QY Sequence 43305, A

Db Sequence 61, SFEOPEARPLFSQAKTLERSVQTKAVNKOLDENISLFLLHSPFLKPAQCKLEWLR

QY Sequence 61 SFEOPEARPLFSQAKTLERSVQTKAVNKOLDENISLFLLHSPFLKPAQCKLEWLR

QY Sequence 61 SFEOPEARPLFSQAKTLERSVQTKAVNKOLDENISLFLLHSPFLKPAQCKLEWLR

QY Sequence 121 FHILYNQDSLIAVLCPHETRIFVRVQLKINNSKIRWMFLIPVKPGVPLAKGTUT

QY Sequence 121 FHILYNQDSLIAVLCPHETRIFVRVQLKINNSKIRWMFLIPVKPGVPLAKGTUT

QY Sequence 121 FHILYNQDSLIAVLCPHETRIFVRVQLKINNSKIRWMFLIPVKPGVPLAKGTUT

QY Sequence 181 HCYKDGLGMDPISLVTKVKVVAEYPPSSAQRLVILLAFTYASTIVSVALVAEDSDNIA

QY Sequence 181 HCYKDGLGMDPISLVTKVKVVAEYPPSSAQRLVILLAFTYASTIVSVALVAEDSDNIA

QY Sequence 241 KLFPIYOKGLKSSLPDYRATYMICQISVKMVENTFNSLASQIKLTKPSLIKQ

QY Sequence 241 KLFPIYOKGLKSSLPDYRATYMICQISVKMVENTFNSLASQIKLTKPSLIKQ

QY Sequence 301 LSCILVLUQRKPEISLGKPPFHLCNVPLDITLHGIBTYDPSPLRMLPHLVSIH

QY Sequence 301 LSCILVLUQRKPEISLGKPPFHLCNVPLDITLHGIBTYDPSPLRMLPHLVSIH

Qy 361 HYTGEETEGMDQOIQYRHLAELITKISLKNNDLHLLASLUREEYSYSSOREMDNKVSL 420
Db 361 HVTGEEETEGMDQOIQYRHLAELITKISLKNNDLHLLASLUREEYSYSSOREMDNKVSL 420
Qy 421 LNEQFLPLIRLJLESKPTRLDVLEBLHEILKEADLUKKOBLFQFVSLSTSGCKYQFLADSD 480
Db 421 LNEQFLPLIRLJLESKPTRLDVLEBLHEILKEADLUKKOBLFQFVSLSTSGCKYQFLADSD 480
Qy 481 TSIMLSLNHPLAPVRILAMNHKKIMKTSKEGVDESFTEKAVALRIGDDNIDVVAISA 540
Db 481 TSIMLSLNHPLAPVRILAMNHKKIMKTSKEGVDESFTEKAVALRIGDDNIDVVAISA 540
Qy 541 FEIPEHFSSEVTINLNLFORAEKSNGEWEVVLKIAADILIKEELS-EN 592
Db 541 FEIPEHFSSEVTINLNLFORAEKSNGEWEVVLKIAADILIKEELS-EN 592
Qy 601 VCLLPFWVINNDTTESEMKAIALYSLSGICSLHPLRGWEBALENVIKSTPKGLIGVA 660
Db 601 VCLLPFWVINNDTTESEMKAIALYSLSGICSLHPLRGWEBALENVIKSTPKGLIGVA 660
Qy 661 NOKMIELADNTNLGDPSMLKVNEDLISVGEESENPLKQVTFHVLTVLVSCLSLKE 720
Db 661 NOKMIELADNTNLGDPSMLKVNEDLISVGEESENPLKQVTFHVLTVLVSCLSLKE 720
Qy 721 THFPAIRVFSLQ 734
Db 721 THFPAIRVFSLQ 734

RESULT 2

US-10-505-928-325

; Sequence 325; Application US/10505928
; Publication No. US20060088532A1
; GENERAL INFORMATION:
; APPLICANT: Ludwig Institute for Cancer Research et al.
; TITLE OF INVENTION: LYMPHATIC ENDOTHELIAL GENES
; FILE REFERENCE: 28967/39178
; CURRENT APPLICATION NUMBER: US/10/505,928
; PRIOR FILING DATE: 2004-08-27
; PRIOR APPLICATION NUMBER: US 60/363, 019
; NUMBER OF SEQ ID NOS: 866
; SOFTWARE: Patentin 3.2
; SEQ ID NO: 325
; LENGTH: 3113
; TYPE: PRT
; ORGANISM: Homo sapiens

US-10-505-928-325

Query Match 2.2%; Score 235; DB 6; Length 3113;
Best Local Similarity 19.0%; Pred. No. 6e-05;
Matches 427; Conservative 357; Mismatches 817; Indels 646; Gaps 101;

Qy 182 CYKDLGPMDIFCISLWTKSVKPAFAEYPCSSAQRLVILAFYASTIVSALVAADDVSDNIAK 241
Db 751 CYQDL-----HAEY---ESLRDL---KSKDASLVTNEDHQSLLA- 785
Qy 242 LPPYIQC-----GLKSSLPDYRAATMICOLISVKUTMFTFVNLASQIKTL 290
Db 786 ---PDDQPMAMHSFANTIGECSMPSPERSE---CRLEADOSPNS----- 824
Qy 291 TKIPSLIKDGSCLLIVLQROKP--BSLGKKPFPHCNVPDLITLHGISF----- 339
Db 825 ---AILQRVDLSELESQKOMMSNLQKQ---CE---BLVQKGIEEENIMKAEQM 873
Qy 340 -----TYDVSPLRIMPLPHLVVSIITHVGTGEEETEGMDGOIYKRHLAELITKISLNL 392
Db 874 QSFVAETSRISKLOEDTSAHONV--VATLSEALENKEKELOLNDKVETEAEQBLKK 932
Qy 393 DHLASLFE----EVYSSQEMOS---NKVSL-----LNEQFLP 427
Db 933 NHILDELSIKEQOLISSETUL- EKKEMSIISLANKREELT-QONGTKEINASLNUQEMN 991

Qy 428 LIRLLESKYPRTDVLEBLHEILKEADLUKKOBLFQFVSLSTSGCKYQFLA-----DSD 480
Db 992 LIQKSES-RANYID-EREKSISELSDQYKQBKLILLORCEBTGNAYEDLSQYKQAQKN 1049
Qy 481 TSIMLSLNHPLAPVRILAMNHKKIMKTSKEGVDESFTEKAVALRIGDDNIDVVAISA 540
Db 1050 SKLECLNB-----CTSLCENRNKNEOLEQKAFAPKEBLTKUAFABERNONML 1100
Qy 536 SAISAFEIPEHFSSEVTINLNLFORAEKSNGEWEVVLKIAADILIKEELS-EN 592
Db 1101 B---LETVOQALASEWTN----QNNKSEAGG-----LKQEINTLKBEQ 1138
Qy 593 DOLSNOVWVCLLPFWVINNDTTESEMKAIALYSLSGICSLHPLRGWEBALENVIKSTPK 652
Db 1139 NMQKEV-----NDLQENBQLMVKMKVTHCONE-----SEPIRNSKERE 1181
Qy 653 PGKLIGVANQ-----KMTELLADNNINGDPSMLKMB-----D 686
Db 1182 SER----NQCNFKPQMDLBVKRISLDNSYNAQLVOLEMARLNKBLKOBSEKEKECLOE 1236
Qy 687 LISV-GEEESFFNLKOKVTHVILSVLVSCLSLKE 720
Db 1237 LQTIRGDLENSNLQDMQSEI---SGLKDCBIDAEVKYTSGPHELTSQDNHQLQSLT 2294
Qy 736 KIKKLESVITAVELPSEWHLEMIDRGPVVELWVHVELNSTORVAVEDSVFLVFSLKK 795
Db 1295 TMNKLNBLKICEIQLAQKYL-----VTLNDSRSECITATRMMAEVOK 1340
Qy 796 PIVALK-----APKSFPKGDIWNNPBQLK-----EDSDYTLHLLJGLFEMM 836
Db 1341 LINNEVKILNDGSLGLHGLYELVEDIPGGEGFQPNQEPVSLAPLDESNSYEHTLSDKEVO 1400
Qy 837 LINGADAVHFRVFLMPLKFIKHLDF-----QLFKFCSTLWITGSSL-----SNPIN 882
Db 1401 M-----HEALQEKPLSLOSEHKTHDHCQMSKMSBLQTYVDSLKAENVLNSTLN 1454
Qy 883 CSVKVLVQTLQYVOCAMLSQKQCKHOLASTS---SPVUTSLINNGSPVKEVRAA 938
Db 1455 FGQDILVKEMQ-LGJUEBGPVPLSSCVPDUSLSSLGDSFYYRALLEOTGD----- 1504
Qy 939 IOCLOCALSGVASPFYLITBHLISKA-----EITSDAVYIQLATLPEELQRE-KK 989
Db 1505 MSLSNLEGAVSANQCSVDBFCSSLQENLTRKETPSPAKGVEEBSLCBVYROSLEK 1564
Qy 990 LKSHQKLSLETLKNUJSCVYSCPSYIAKDMVKGUVGNGR---MVLSQLPMAEOLKE 1044
Db 1565 LEERKMEQSGIMKIN-----KEIQLEOLISSERBQBLDCLRKOYLSENEQWOOK 1611
Qy 1045 I-----QKEPTAVLKOBAMVUHLTGKYNEPSSVSLNEDP-----KSJD 1083
Db 1612 LTSVTELEMESKLAEEKKOTBOLSLBLEVARLQLOGDLSRSLGIDTEDAIORNESC 1671
Qy 1084 IFIKAVHTTKELYAGMPTIQTIALEKITKAPPFAISDEKVKQULRMLFDLLVCKNSC 1143
Db 1672 --ISKEHTSB-----TTERTPKHQHICDKDQAQDQI-----NLDIEKI 1708
Qy 1144 AQTVSASVFKISVNAEQVRIELEPPDCAKPLGTWQKRRQMQQKSQDLESVOEVGGSY 1203
Db 1709 TETGALKPTGECSGEQSPDTNYPGDXQGS-----SECUSLSEFS- 1751
Qy 1204 WQRVTLILELQHKKKLRSQPLQVPTLNF---IL-----SRQLEPLPQEQGNME--Y 1250
Db 1752 GPNALVPMDFLGNQDQIHNLQRLVYKETSNNENRLUHVBDRKVESLNEMKELDSKUH 1811
Qy 1251 TKQLLSCLINICOLSPROGKPKDILBEEKFNVELIVQCTIRSEMPOTHHALLGT 1310
Db 1812 LOEVOLMLTKTACIELEKIVGELKKENDSE-----KUBYFSCDHQELJORVET 1861
Qy 1311 VAGJFPDKVHNIMSIFTPGANWRLLDTYSFOVINKVVMFALQOSDG---DSI 1366
Db 1862 SEGUNSLDEMHADKSREPDIQDNTVAKUNDWSKERFLDVNETE---SRIRSEKASTERBAL 1918
Qy 1367 EVSRNVEITV-----K11SVFVDALPHVPERRLPILVQI-----VDTLGAK 1410